

ORIGINAL ARTICLE

Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis

Peter C. Taylor, M.D., Ph.D., Edward C. Keystone, M.D.,
 Désirée van der Heijde, M.D., Ph.D., Michael E. Weinblatt, M.D.,
 Liliana del Carmen Morales, M.D., Jaime Reyes Gonzaga, M.D.,
 Sergey Yakushin, M.D., Taeko Ishii, M.D., Kahaku Emoto, M.D.,
 Scott Beattie, Ph.D., Vipin Arora, Ph.D., Carol Gaich, Pharm.D.,
 Terence Rooney, M.D., Douglas Schlichting, R.N., Ph.D.,
 William L. Macias, M.D., Ph.D., Stephanie de Bono, M.D., Ph.D.,
 and Yoshiya Tanaka, M.D., Ph.D.

ABSTRACT

BACKGROUND

Baricitinib is an oral, reversible inhibitor of the Janus kinases JAK1 and JAK2 that may have therapeutic value in patients with rheumatoid arthritis.

METHODS

We conducted a 52-week, phase 3, double-blind, placebo- and active-controlled trial in which 1307 patients with active rheumatoid arthritis who were receiving background therapy with methotrexate were randomly assigned to one of three regimens in a 3:3:2 ratio: placebo (switched to baricitinib after 24 weeks), 4 mg of baricitinib once daily, or 40 mg of adalimumab (an anti-tumor necrosis factor α monoclonal antibody) every other week. End-point measures evaluated after adjustment for multiplicity included 20% improvement according to the criteria of the American College of Rheumatology (ACR20 response) (the primary end point), the Disease Activity Score for 28 joints (DAS28), the Health Assessment Questionnaire–Disability Index, and the Simplified Disease Activity Index at week 12, as well as radiographic progression of joint damage as measured by the van der Heijde modification of the total Sharp score (mTSS) (range, 0 to 448, with higher scores indicating greater structural joint damage) at week 24.

RESULTS

More patients had an ACR20 response at week 12 with baricitinib than with placebo (primary end point, 70% vs. 40%, $P<0.001$). All major secondary objectives were met, including inhibition of radiographic progression of joint damage, according to the mTSS at week 24 with baricitinib versus placebo (mean change from baseline, 0.41 vs. 0.90; $P<0.001$) and an increased ACR20 response rate at week 12 with baricitinib versus adalimumab (70% vs. 61%, $P=0.014$). Adverse events, including infections, were more frequent through week 24 with baricitinib and adalimumab than with placebo. Cancers were reported in five patients (two who received baricitinib and three who received placebo). Baricitinib was associated with reductions in neutrophil counts and increases in levels of creatinine and low-density lipoprotein cholesterol.

CONCLUSIONS

In patients with rheumatoid arthritis who had had an inadequate response to methotrexate, baricitinib was associated with significant clinical improvements as compared with placebo and adalimumab. (Funded by Eli Lilly and Incyte; ClinicalTrials.gov number, NCT01710358.)

From the Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences and the Kennedy Institute of Rheumatology, University of Oxford, Oxford, United Kingdom (P.C.T.); the Rebecca MacDonald Centre, Mount Sinai Hospital, University of Toronto, Toronto (E.C.K.); Leiden University Medical Center, Leiden, the Netherlands (D.H.); Brigham and Women's Hospital, Boston (M.E.W.); Instituto Reumatológico Strusberg, Córdoba, Argentina (L.C.M.); Centro de Investigación Clínica Especializada, Mexico City (J.R.G.); Ryazan Regional Clinical Cardiology Dispensary, Ryazan, Russia (S.Y.); Eli Lilly, Indianapolis (T.L., S. Beattie, V.A., C.G., T.R., D.S., W.L.M., S. de Bono); and AstraZeneca K.K., Osaka (K.E.), and the First Department of Internal Medicine, School of Medicine, University of Occupational and Environmental Health, Japan, Kitakyushu (Y.T.) — both in Japan. Address reprint requests to Dr. Taylor at the Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, Kennedy Institute of Rheumatology, University of Oxford, Windmill Rd., Headington, Oxford OX3 7LD, United Kingdom, or at peter.taylor@kennedy.ox.ac.uk.

N Engl J Med 2017;376:652-62.
 DOI: 10.1056/NEJMoa1608345
 Copyright © 2017 Massachusetts Medical Society.

RHEUMATOID ARTHRITIS IS A SYSTEMIC autoimmune disease characterized by inflammatory synovitis and progressive joint destruction, which are associated with severe disability and increased mortality. Progress in treatment with the use of conventional synthetic disease-modifying antirheumatic drugs (DMARDs), such as methotrexate, and biologic DMARDs that target tumor necrosis factor (TNF) has made clinical remission a realistic target.¹

Activated Janus kinases (JAKs) play pivotal roles in intracellular signaling from cell-surface receptors for multiple cytokines implicated in the pathologic processes of rheumatoid arthritis.² Baricitinib, an orally available small molecule, provides reversible inhibition of JAK1 and JAK2 and has shown clinical efficacy in studies involving patients with rheumatoid arthritis.³⁻⁶

The RA-BEAM trial was a global, phase 3, double-blind, placebo- and active-controlled trial involving patients with moderately to severely active rheumatoid arthritis. The trial was designed to assess changes in disease activity, structural preservation, and patient-reported outcomes (including physical function), in addition to the safety and side-effect profile of a regimen of 4 mg of oral baricitinib once daily, in patients who had had an inadequate response to methotrexate. Comparisons were made with placebo and the TNF- α inhibitor adalimumab, a standard-of-care biologic DMARD for patients with moderately to severely active rheumatoid arthritis.

METHODS

PATIENTS

Patients were 18 years of age or older and had active rheumatoid arthritis (≥ 6 tender joints of 68 examined, ≥ 6 swollen joints of 66 examined, and a high-sensitivity serum C-reactive protein level of ≥ 6 mg per liter). Patients had had an inadequate response to methotrexate, having received 12 weeks or more of therapy before trial entry, including 8 weeks or more at stable doses of 15 to 25 mg per week, unless lower doses were clinically indicated. At baseline, patients were required to have either three or more joint erosions (diagnosed on the basis of centrally evaluated radiographs of hands, wrists, and feet) or one or more joint erosions plus seropositivity for rheumatoid factor or anti-citrullinated peptide antibodies. The criteria for exclusion includ-

ed previous biologic DMARD therapy, selected laboratory abnormalities, and recent clinically serious infection. Patients with evidence of latent tuberculosis could enroll if appropriate treatment had commenced 4 weeks or more before randomization.

STUDY DESIGN AND OVERSIGHT

RA-BEAM was a randomized, double-blind, placebo- and active-controlled, parallel-group trial that was conducted for 52 weeks at 281 centers in 26 countries. Patients were randomly assigned in a ratio of 3:3:2 to receive placebo, 4 mg of baricitinib once daily, or 40 mg of subcutaneous adalimumab every other week, in addition to existing background therapy (including methotrexate). At week 24, patients receiving placebo were switched to baricitinib and were unaware of the change in treatment. Patients with an estimated glomerular filtration rate of 40 to less than 60 ml per minute per 1.73 m² (approximately 4%) received 2 mg of baricitinib if assigned to baricitinib treatment. Concomitant stable doses of conventional synthetic DMARDs, nonsteroidal antiinflammatory drugs, analgesics, or glucocorticoids (≤ 10 mg of prednisone or the equivalent per day) were permitted.

At week 16, patients whose counts of tender and swollen joints were reduced by less than 20% from baseline at both week 14 and week 16 received open-label rescue treatment (4 mg of baricitinib). Afterward, patients received rescue treatment at investigators' discretion on the basis of joint counts. Patients who completed the trial were eligible to enter a long-term extension study or were seen for follow-up (up to approximately 28 days after the end of treatment).

The trial was designed by the sponsor, Eli Lilly, an academic advisory board that included authors who were not employees of Eli Lilly, and Incyte. The study was conducted in accordance with ethical principles of the Declaration of Helsinki and Good Clinical Practice guidelines and approved by each center's institutional review board or ethics committee. All the patients provided written informed consent. The trial commenced in November 2012 and was completed in September 2015 and enrolled patients from November 2012 through September 2014. Eli Lilly or its representatives provided data, laboratory, and site-monitoring services. Adalimumab was manufactured by AbbVie and purchased

through commercial sources. All the authors participated in the analysis and interpretation of the data, reviewed the draft and the final manuscript, provided critical comment, and made the decision to submit the manuscript for publication. The authors vouch for the veracity and completeness of the data and analyses and for the fidelity of this report to the protocol.

EFFICACY

For the primary end point, baricitinib was compared with placebo on the basis of the proportion of patients at week 12 with a 20% response according to the criteria of the American College of Rheumatology (ACR20 response).⁷ The ACR20 response is a reduction of 20% or more in the number of tender and swollen joints and an improvement of 20% or more in at least three of the following ACR core measures: a patient's assessment of pain, a physician's global assessment of disease, a patient's global assessment of disease, physical function as assessed by the Health Assessment Questionnaire–Disability Index (HAQ-DI), and the level of acute-phase reactant (see Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org, for a description of this and other measures of efficacy).

A major secondary end point was the progression of joint damage from baseline to week 24 detected on radiography as assessed with the use of the van der Heijde modification of the total Sharp score (mTSS; range, 0 to 448, with higher scores indicating greater structural joint damage). Radiographs were scored by two readers who were unaware of the chronologic order in which the radiographs were obtained, patient identity, and group assignment, with the average score between readers used for analysis.^{8,9} Other major secondary end points (evaluated at week 12) included changes in physical function, as assessed with the HAQ-DI (range 0 to 3, with higher scores indicating greater disability),^{10,11} and in disease activity, as assessed with the Disease Activity Score for 28 joints (DAS28) with the use of high-sensitivity C-reactive protein (DAS28-CRP), with higher scores indicating greater disease activity. Major secondary end points also included remission rate as measured with the Simplified Disease Activity Index (SDAI), with disease remission defined as a score of ≤ 3.3 (range, 0.1 to 86.0, with higher scores

indicating greater disease activity), and patient-reported outcomes as recorded daily in an electronic diary, including morning joint stiffness (measured in minutes and in severity), degree of tiredness, and degree of joint pain (with severity of morning joint stiffness, tiredness, and joint pain measured on a numeric rating scale of 0 to 10, with higher values indicating worse ratings).

Comparisons that were controlled for multiplicity included baricitinib versus placebo with respect to all major secondary end points and baricitinib versus adalimumab at week 12 for the ACR20 response and the change from baseline in DAS28-CRP. Secondary and exploratory end points that were not controlled for multiplicity involved comparisons of all applicable groups at each time point (with no use of placebo after week 24) for all efficacy measures mentioned above and other measures. These other measures included ACR50 and ACR70 response rates (i.e., 50% and 70% improvement, respectively), DAS28 with the use of the erythrocyte sedimentation rate (DAS28-ESR), and the Clinical Disease Activity Index.¹²⁻¹⁶

SAFETY

Clinical laboratory tests, vital signs, and other safety assessments were performed at scheduled visits. The incidence and severity of all adverse events were recorded. The National Institutes of Health Common Terminology Criteria for Adverse Events (version 3.0) and the National Cholesterol Education Program categories were used to describe laboratory abnormalities. An independent data and safety monitoring committee regularly reviewed data from this and contemporaneous phase 3 studies of baricitinib. An independent clinical end-point committee adjudicated potential cardiovascular events.

STATISTICAL ANALYSIS

We estimated that an unbalanced randomization of approximately 1280 patients (480 assigned to placebo, 480 to baricitinib, and 320 to adalimumab) would provide sufficient power for comparisons of the ACR20 response rates at week 12 between baricitinib and placebo (estimated power for test of superiority, >95%) and with adalimumab (estimated power for test of noninferiority, 93%), assuming rates of 35% with placebo and 60% with both baricitinib and adalimumab. A prespecified, closed, sequentially rejective, weight-

ed Bonferroni multiple-testing procedure was used to control for type I errors related to the primary and major secondary objectives^{17,18} (Fig. S2 in the Supplementary Appendix), including two assessments against adalimumab: a test of superiority with respect to DAS28-CRP and a test of noninferiority with respect to ACR20. A prespecified noninferiority margin of 12% was chosen on the basis of its use in previous head-to-head trials involving rheumatoid arthritis^{19,20} and of a Bayesian meta-analysis of multinational, placebo-controlled trials involving similar populations, which determined that a 12% margin would be consistent with the natural variability in reported ACR20 response rates. In the plan for multiple comparisons, if noninferiority was shown, the superiority of baricitinib to adalimumab would be evaluated. After rejection of the primary null hypothesis for the comparison of baricitinib with placebo on the basis of ACR20, the type I error rate was allocated among major secondary hypotheses according to the multiple comparisons procedure. At each step of the procedure, rejection of any null hypothesis resulted in reallocation of the error rate assigned to that hypothesis among the remaining hypotheses until no further hypotheses could be rejected or all were rejected.

The modified intention-to-treat efficacy-analysis set included all the patients who had undergone randomization and been treated with at least one dose of the study drug. Comparisons of categorical efficacy end points were made with the use of logistic regression, with region, baseline joint-erosion status (per entry criteria), and randomized treatment group included in the model. Treatment comparisons of continuous efficacy end points were made with the use of analysis of covariance (ANCOVA), with adjustments for treatment, region, joint-erosion status at baseline, and baseline value. Fisher's exact test was used for the analysis of other categorical data. Continuous safety data were analyzed with the use of ANCOVA, with adjustment for baseline value and treatment. Analyses were assessed at a two-sided alpha level of 0.05 unless otherwise defined in the multiple testing procedure (Fig. S2 in the Supplementary Appendix).

Patients who received rescue treatment or who discontinued the study treatment were therefore considered not to have had a response (non-responder imputation) for all categorical efficacy

outcomes. These patients also had their last observations before rescue treatment or study discontinuation carried forward (modified last observation carried forward) for analyses of continuous efficacy data. Regarding the assessment of HAQ-DI and DAS28 as major end points, if the reason for discontinuation was an adverse event, the baseline observation was substituted at the primary analysis time point (modified baseline observation carried forward). For radiographic measures, scores at week 24 or 52 that were missing or obtained subsequent to rescue treatment or a planned switch to baricitinib as defined in the protocol (available at NEJM.org) were imputed with the use of linear extrapolation from baseline and the most recent post-baseline data obtained before or at the initiation of rescue or switch therapy. In a supportive analysis of radiographic measures, all available observed data (including data obtained after rescue or switch therapy) were used, with missing data imputed by means of the last observation carried forward and patients analyzed according to the group to which they were originally assigned. Alternative methods of analysis (e.g., mixed models for repeated measures and tipping-point analyses) were conducted to ensure that conclusions were robust and not dependent on mechanisms used to account for missing data. Safety data were analyzed according to the initially assigned group until the initiation of rescue or switch therapy or the completion of the treatment period.

RESULTS

PATIENTS

Among 2949 screened patients, 1307 underwent randomization and 1305 were treated and qualified for analysis (Fig. S1 in the Supplementary Appendix). The most common reasons why patients did not proceed to randomization were a high-sensitivity CRP level of less than 6 mg per liter and the absence of joint erosions. Baseline demographic and clinical characteristics were similar among the groups (Table 1, and Table S2 in the Supplementary Appendix). Most patients (>99%) were receiving background methotrexate; the majority had previously received at least two conventional synthetic DMARDs. Rescue rates for the placebo, baricitinib, and adalimumab groups were 27%, 9%, and 15%, respectively

Characteristic	Placebo (N=488)	Baricitinib (N=487)	Adalimumab (N=330)
Age — yr	53±2	54±2	53±12
Female sex — no. (%)	382 (78)	375 (77)	251 (76)
Duration of rheumatoid arthritis — yr	10±9	10±9	10±9
Positive for anti-cyclic citrullinated peptide — no. (%)†	424 (87)	427 (88)	295 (89)
Positive for rheumatoid factor — no. (%)‡	451 (92)	439 (90)	301 (91)
Had ≥3 erosions — no./total no. (%)	371/488 (76)	371/487 (76)	245/327 (75)
mTSS units	45±50	43±50	44±51
Erosion score	27±29	25±28	26±29
Score for narrowing of joint space	18±23	17±23	18±24
Swollen-joint count, of 66 joints examined	16±9	15±8	15±9
Tender-joint count, of 68 joints examined	23±14	23±13	23±14
Scores for global and pain assessment§			
Physician's Global Assessment	64±17	66±17	65±17
Patient's Global Assessment	61±23	63±21	64±21
Patient's assessment of pain	60±23	62±22	61±23
HAQ-DI	1.55±0.67	1.57±0.68	1.59±0.70
High-sensitivity C-reactive protein — mg/liter¶	20±21	22±23	22±21
ESR — mm/hr	49±26	49±26	48±26
DAS28-CRP	5.7±1.0	5.8±0.9	5.8±0.9
DAS28-ESR	6.4±1.0	6.5±0.9	6.4±1.0
Simplified Disease Activity Index	40±13	40±13	40±13

* Plus-minus values are means ±SD. The total number of patients in each group is the number of patients who were randomly assigned to that group and who received at least one dose of the assigned study medication. There were no clinically significant between-group differences at baseline. DAS28 denotes the 28-joint Disease Activity Score, which is based on the C-reactive protein level (DAS28-CRP) or on the erythrocyte sedimentation rate (DAS28-ESR); HAQ-DI Health Assessment Questionnaire-Disability Index, in which the range of scores is 0 to 3, with higher scores indicating greater disability; and mTSS the van der Heijde modification of the total Sharp score, in which the range is 0 to 448, with higher scores indicating greater damage. See Table S2 in the Supplementary Appendix for information on baseline characteristics and disease activity according to geographic region.

† Positivity for anti-cyclic citrullinated peptide antibody was defined as a value that exceeded the upper limit of the normal range (i.e., >10 U per milliliter).

‡ Positivity for rheumatoid factor was defined as a value that exceeded the upper limit of the normal range (i.e., >14 IU per milliliter).

§ Values for the Physician's Global Assessment, the Patient's Global Assessment, and the patient's assessment of pain range from 0 to 100 mm (visual analogue scale), with higher values indicating greater levels of disease activity or pain.

¶ The upper limit of the normal range for the level of high-sensitivity C-reactive protein is 3.0 mg per liter.

|| In the Simplified Disease Activity Index, scores range from 0.1 to 86.0, with higher scores indicating greater disease activity and a score of 3.3 or less indicating remission.

(Fig. S1 in the Supplementary Appendix). Approximately 86% of the patients participated in the trial for 52 weeks, and a large majority entered a long-term extension study (Fig. S1 in the Supplementary Appendix).

EFFICACY

At week 12, the primary ACR20 response rate for baricitinib was 70% as compared with 40% for placebo ($P<0.001$). Significant improvements with baricitinib as compared with placebo were seen

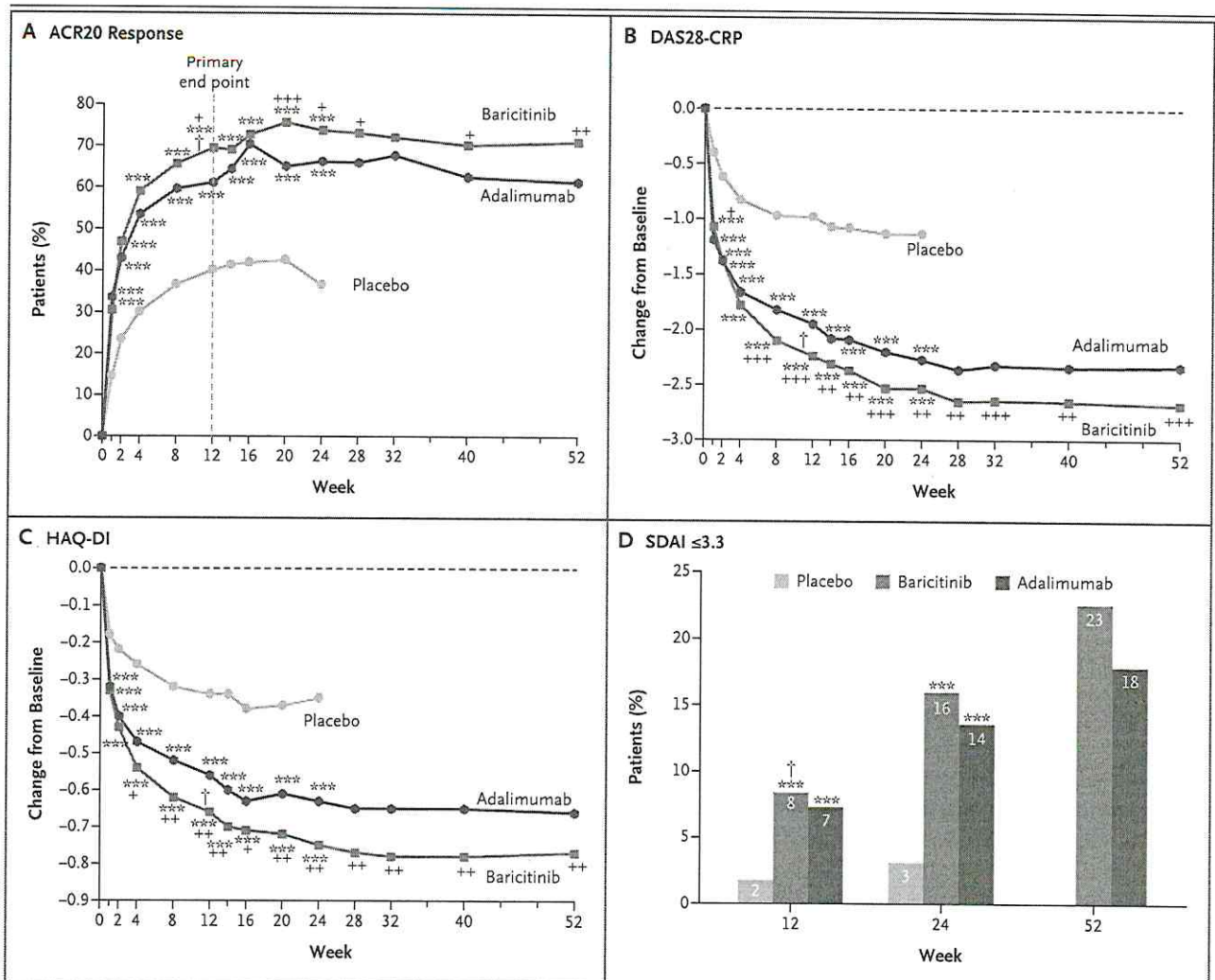


Figure 1. Primary and Secondary Efficacy End Points.

Panel A shows the percentage of patients who had 20% improvement according to the criteria of the American College of Rheumatology (ACR20 response). The vertical line at 12 weeks indicates the primary efficacy time point. Panel B shows the least squares mean (LSM) change from baseline in the 28-joint Disease Activity Score, based on the level of high-sensitivity C-reactive protein (DAS28-CRP). For data that were missing because of receipt of rescue therapy or premature discontinuation of the study or study treatment, a modified last-observation-carried-forward (mLOCF) method was used to incorporate the last observation before rescue or discontinuation. Panel C shows the LSM change from baseline in the Health Assessment Questionnaire–Disability Index (HAQ-DI) score, calculated with the use of the mLOCF method; negative changes from baseline indicate improvement. Panel D shows the percentage of patients with a Simplified Disease Activity Index (SDAI) score of 3.3 or less (scores range from 0.1 to 86.0, with higher scores indicating greater disease activity and a score of 3.3 or less indicating remission) at weeks 12, 24, and 52. The DAS28-CRP and HAQ-DI were analyzed at week 12 by means of the modified baseline-observation-carried-forward method, which substituted the baseline observation for patients who had left the study because of an adverse event. Panels A through D include all the patients in the modified intention-to-treat efficacy analysis set (i.e., all the patients who underwent randomization and were treated), which included 1305 patients. Three asterisks denote $P < 0.001$ for supportive analyses comparing baricitinib at the 4-mg dose or adalimumab with placebo, without adjustment for multiple comparisons. One plus sign denotes $P \leq 0.05$, two plus signs $P \leq 0.01$, and three plus signs $P < 0.001$ for supportive analyses comparing 4 mg of baricitinib with adalimumab, without adjustment for multiple comparisons. A dagger denotes comparisons of baricitinib with placebo and baricitinib with adalimumab for the primary and key secondary end points that are statistically significant as calculated with the graphical method for multiple testing, with the studywise error rate strongly controlled at an alpha level of 0.05 for multiple comparisons.

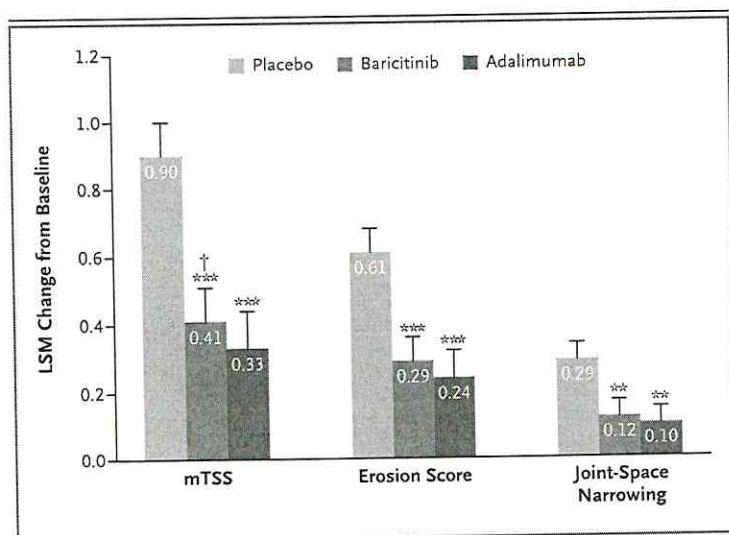


Figure 2. Inhibition of Radiographic Progression of Structural Joint Damage at Week 24.

The least squares mean (LSM) change from baseline in structural progression was evaluated with the use of the van der Heijde modification of the total Sharp score (mTSS), with scores ranging from 0 to 448, with higher scores indicating greater structural joint damage. Also shown are scores for erosion and joint-space narrowing. This evaluation included all the patients with a baseline measurement and at least one postbaseline radiograph for the assessment of progression of structural joint damage, which totaled 1234 patients at week 24. Scores that were missing or acquired subsequent to rescue treatment or to a treatment switch as defined in the protocol were imputed with the use of linear extrapolation from baseline and the most recent postbaseline data obtained before or at the initiation of rescue or switch therapy. T bars denote the standard error. Two asterisks denote $P \leq 0.01$ and three asterisks $P < 0.001$ for 4 mg of baricitinib or adalimumab versus placebo. A dagger indicates that comparisons between baricitinib and placebo and between baricitinib and adalimumab with respect to the primary and key secondary end points are statistically significant as calculated with the graphical method for multiple testing, with the studywise error rate strongly controlled at an alpha level of 0.05 for multiple comparisons.

at week 12 with respect to all major secondary end points in the analyses that were controlled for multiple testing, including HAQ-DI, DAS28-CRP, remission according to the SDAI (Fig. 1), and daily diary measures (i.e., duration and severity of morning joint stiffness, worst tiredness, and worst joint pain) (Fig. S5 in the Supplementary Appendix). A significant reduction in radiographic progression of structural joint damage was seen at week 24 for both baricitinib and adalimumab as compared with placebo (Fig. 2).

Baricitinib was found to be noninferior to adalimumab at week 12 for the ACR20 response, with a noninferiority margin of 12% (70% for baricitinib and 61% for adalimumab; 95% confidence interval for the difference between groups,

2% to 15%). According to the statistical analysis plan, baricitinib was therefore considered to be significantly superior to adalimumab ($P = 0.01$). In addition, baricitinib was superior to adalimumab according to the mean change in DAS28-CRP at week 12 (-2.24 for baricitinib vs. -1.95 for adalimumab, $P < 0.001$).

Results for the comparison of baricitinib with placebo and with adalimumab for other secondary and exploratory end points, including ACR50 and ACR70 response rates, DAS28-ESR, SDAI, Clinical Disease Activity Index, and individual components of the ACR response rate, are provided in the Supplementary Appendix. Significant improvements in many efficacy measures, including the ACR20 response rate, DAS28-CRP, and patient-reported outcomes, were observed as early as week 1 for baricitinib as compared with placebo and as early as weeks 2 through 4 for baricitinib as compared with adalimumab (Fig. 1, and Fig. S5 in the Supplementary Appendix). Measures of efficacy were maintained or improved through week 52. In prespecified supportive analyses that were based on linear extrapolation or that used all available observed data, including data obtained after rescue or switch therapy (e.g., from placebo to baricitinib after week 24), a significant reduction in radiographic progression as compared with placebo was observed for both baricitinib and adalimumab at weeks 24 and 52 (Figs. S6 and S7 in the Supplementary Appendix).

SAFETY

Rates of discontinuation resulting from adverse events from baseline through week 24 were 3% with placebo, 5% with baricitinib, and 2% with adalimumab (Table 2). Rates of serious adverse events through week 24 were 5% with placebo, 5% with baricitinib, and 2% with adalimumab, with no particular type of event contributing to the lower rate observed with adalimumab. Five deaths were reported: one in the placebo group, two in the baricitinib group, one in the adalimumab group, and one in a patient in the placebo group who received rescue treatment with baricitinib (Fig. S1 in the Supplementary Appendix). Rates of cancer through 24 weeks were less than 1% each with placebo and baricitinib and 0 with adalimumab (Table 2).

Adverse events were more frequent with baricitinib and adalimumab than with placebo

Table 2. Safety and Laboratory Data, Week 0 to Week 24 and Week 0 to Week 52.*

Variable	Week 0 to Week 24		Week 0 to Week 52	
	Placebo (N = 488)	Baricitinib (N = 487)	Adalimumab (N = 330)	Baricitinib (N = 487) / Adalimumab (N = 330)
Safety data				
Treatment exposure — patient-yr	197.7	215.0	141.9	430.7
Serious adverse events — no. (%)†	22 (5)	23 (5)	6 (2)	38 (8)
Any adverse event after start of therapy — no. (%)	295 (60)	347 (71)	224 (68)	384 (79)
Withdrawal because of adverse event — no. (%)	17 (3)	24 (5)	7 (2)	36 (7)
Infection — no. (%)	134 (27)	176 (36)	110 (33)	233 (48)
Herpes zoster — no. (%)	2 (<1)	7 (1)	4 (1)	11 (2)
Tuberculosis — no. (%)	0	0	1 (<1)	0
Serious infection — no. (%)	7 (1)	5 (1)	2 (<1)	10 (2)
Cancer — no. (%)	3 (<1)	2 (<1)	0	3 (<1)
Nonmelanoma skin cancer — no. (%)	1 (<1)	0	0	0
Breast cancer — no. (%)	1 (<1)	1 (<1)	0	1 (<1)
Squamous-cell cancer — no. (%)	0	1 (<1)	0	1 (<1)
Ovarian cancer — no. (%)	1 (<1)	0	0	0
Clear-cell renal carcinoma — no. (%)	0	0	0	0
Major adverse cardiovascular event — no. (%)‡	0	1 (<1)	0	1 (<1)
Gastrointestinal perforation — no. (%)	0	0	0	2 (<1)
Laboratory data — least-squares-mean change from baseline§				
Hemoglobin — g/dl	0.09±0.05	0.09±0.05	0.53±0.06¶	0.19±0.05
Neutrophils — per mm ³	-300±100	-1040±90¶	-1400±110¶	-1230±90
Lymphocytes — per mm ³	-20±30	20±30	330±40¶	-30±30
Platelet counts — per mm ³	-8000±3000	9000±3000¶	-43,000±3000¶	13,000±3000
Alanine aminotransferase — IU/liter	1.0±1.0	5.9±0.9¶	5.3±1.1¶	5.5±0.7
Creatinine — mg/dl	0.011±0.006	0.071±0.005¶	0.039±0.006¶	0.086±0.005
Creatine phosphokinase — IU/liter	5±4	54±3¶	22±4¶	67±4
Cholesterol — mg/dl				
LDL	-3±1	16±1¶	7±1¶	18±1
HDL	-0.1±0.6	9.5±0.6¶	4.8±0.7¶	8.0±0.6

* Plus-minus values are means ±SE. HDL denotes high-density lipoprotein, and LDL low-density lipoprotein. To convert the values for creatinine to micromoles per liter, multiply by 88.4. To convert the values for cholesterol to millimoles per liter, multiply by 0.02586.

† Serious adverse events are reported on the basis of conventional International Conference on Harmonisation definitions and not on the basis of the protocol; the protocol required that adverse events or laboratory abnormalities leading to permanent discontinuation of the study drug be designated as serious adverse events. The data shown are numbers and percentages of patients with serious adverse events, up to the time of rescue therapy.

‡ A major adverse cardiovascular event was defined as cardiovascular death, myocardial infarction, or stroke, as adjudicated by an independent cardiovascular evaluation committee.

§ Laboratory values are reported as the least-squares-mean change from baseline at week 24 and week 52.

¶ P<0.001 for the comparison between baricitinib and placebo on the basis of analysis of covariance (ANCOVA).

|| P≤0.01 for the comparison between adalimumab and placebo on the basis of ANCOVA.

(Table 2); rates of adverse events (including infections) associated with baricitinib and adalimumab were similar through 52 weeks. Rates of serious infection were similar with placebo, baricitinib, and adalimumab through week 24 (1%, 1%, and <1%, respectively) and with baricitinib and adalimumab through week 52 (2% each). Herpes zoster was seen in all groups (at a rate of 2% in both the baricitinib and adalimumab groups); most cases occurred in Asia. One patient in the baricitinib group had a herpes zoster rash that was distributed beyond the primary and adjacent dermatomes, but the patient recovered without complications.

Mean changes in laboratory values from baseline, increases in grade according to the National Institutes of Health Common Terminology Criteria for Adverse Events, and observed mean values for selected laboratory analytes through weeks 24 and 52 are provided in Table 2, and in Tables S7 and S11 in the Supplementary Appendix. With baricitinib, the mean hemoglobin level was stable at week 24 and increased from baseline at week 52; with adalimumab, the mean level increased from baseline at both week 24 and week 52. Reductions in neutrophil counts were observed with baricitinib and adalimumab. Early increases in lymphocyte counts were seen with baricitinib and adalimumab but not with placebo; counts returned to baseline through weeks 24 and 52 in the baricitinib group, and counts remained elevated in the adalimumab group. Modest increases in platelet counts were seen with baricitinib, whereas a decrease was seen with adalimumab. There was no significant difference between groups in rates of thrombocytosis as defined in the protocol (>600,000 cells per cubic millimeter).

Increases in alanine aminotransferase levels were observed with baricitinib and adalimumab (mean changes from baseline, 5.9 and 5.3 IU per liter, respectively, at week 24); most increases were transient, and no elevations of grade 2 or higher coincided with increases in bilirubin levels. A total of five patients (three in the baricitinib group and two in the adalimumab group) permanently discontinued treatment for adverse events related to the liver by week 52. Small increases in the mean serum creatinine level were seen with baricitinib and adalimumab; most increases in grade were transient, and none ex-

ceeded grade 1 in patients taking baricitinib or grade 2 in patients taking adalimumab. Serum creatine phosphokinase levels increased with baricitinib and adalimumab. Among the few patients with elevations at grade 3 or 4, most had associated increases in physical activity, abnormally high values at baseline, or both (Table S12 in Supplementary Appendix). Levels of low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol increased with baricitinib and adalimumab as compared with placebo at week 24, with LDL cholesterol levels remaining stable and HDL cholesterol levels decreasing slightly (but increasing relative to baseline values) between weeks 24 and 52 in the two groups. For some analytes (e.g., serum levels of creatinine, creatine phosphokinase, LDL cholesterol, and HDL cholesterol), changes were directionally similar with baricitinib and adalimumab but larger with baricitinib.

DISCUSSION

In this study of patients who had had an inadequate response to methotrexate and had not been treated with biologic DMARDs, a regimen of 4 mg of baricitinib once daily was compared with placebo or 40 mg of adalimumab every other week. All patients received background treatment with methotrexate, a context in which adalimumab has proved to be most efficacious.²⁰ All objectives that were included within the context of strong multiplicity control were met. Baricitinib showed significant clinical benefits as compared with placebo at week 12 along with greater efficacy than adalimumab according to ACR20 response rate and diminished disease activity according to DAS28-CRP. As compared with placebo, significant inhibition of radiographic progression of disease was observed at week 24 with both baricitinib and adalimumab.

Adverse events, including infections, were more frequent with baricitinib and adalimumab than with placebo through week 24. Serious adverse events through week 24 were more frequent with baricitinib and placebo than with adalimumab. Baricitinib and adalimumab were both associated with reductions in neutrophil counts, increases from baseline in aminotransferase and creatinine levels, and increases from baseline in LDL and HDL cholesterol levels.

This study has several limitations. Although the design allowed for the use of placebo for 24 weeks, a rescue option was mandated at 16 weeks for patients who did not show a response to treatment to address ethical concerns about continuing placebo in patients with active disease. Although 27% of the patients who had been assigned to placebo received rescue treatment and 11% of those assigned to placebo discontinued placebo before receiving rescue treatment or being switched to baricitinib at week 24 (in keeping with the study protocol), patients remained unaware of their initially assigned treatment until the end of the study. The study enrolled patients who had had an inadequate response to methotrexate. Among these patients, only 15 to 18% in each treatment group were receiving other conventional synthetic DMARDs. Thus, the study has a limited capacity to assess the effectiveness of baricitinib when used in combination with conventional synthetic DMARDs other than methotrexate.

Despite advances in the management of rheumatoid arthritis, limitations in treatment remain. These include limitations associated with the parenteral delivery of biologic drugs, the fact that not all patients have a response to conven-

tional synthetic DMARDs or biologic DMARDs (whether administered alone or in combination), and the frequent need for polypharmacy. Furthermore, conventional synthetic DMARDs have a relatively slow onset of action and are generally less effective than biologic DMARDs in inhibiting structural joint damage.²¹⁻²⁴ Our study showed that for the outcome measure used as the primary end point, the combination of baricitinib plus methotrexate was superior to adalimumab plus methotrexate, the latter being a current standard-of-care treatment in this patient population.

In conclusion, in patients with active rheumatoid arthritis despite receiving therapy with methotrexate, the addition of once-daily oral baricitinib was associated with improvements in signs and symptoms, physical function, patient-reported outcomes, and progression of structural joint damage as compared with placebo and with improvements in ACR20 response and DAS28-CRP as compared with adalimumab.

Supported by Eli Lilly and Incyte.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Stephanie Colvin, Ph.D., at Eli Lilly for assistance with tables and figures and with manuscript preparation and process support.

REFERENCES

- Smolen JS, Aletaha D, Koeller M, Weisman MH, Emery P. New therapies for treatment of rheumatoid arthritis. *Lancet* 2007;370:1861-74.
- O'Shea JJ, Kontzias A, Yamaoka K, Tanaka Y, Laurence A. Janus kinase inhibitors in autoimmune diseases. *Ann Rheum Dis* 2013;72:Suppl 2:iii111-iii115.
- 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 on background methotrexate therapy: a 12-week double-blind, randomized placebo-controlled study. *J Rheumatol* 2016;43:504-11.
- Genovese MC, Kremer J, Zamani O, et al. Baricitinib in patients with refractory rheumatoid arthritis. *N Engl J Med* 2016; 374:1243-52.
- LB0001 baricitinib, an oral Janus kinase (JAK)1/JAK2 inhibitor, in patients with active rheumatoid arthritis (RA) and an inadequate response to CDMArD therapy: results of the phase 3 RA-BUILD Study. *Ann Rheum Dis* 2015 (http://ard.bmj.com/content/74/Suppl_2/79.2).
- Felson DT, Anderson JJ, Boers M, et al. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727-35.
- van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000;27:261-3.
- van der Heijde DM. Assessment of radiographs in longitudinal observational studies. *J Rheumatol Suppl* 2004;69:46-7.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
- Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789-93.
- Smolen JS, Breedveld FC, Schiff MH, et al. A Simplified Disease Activity Index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003; 42:244-57.
- 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:Suppl 39:S100-S108.
- Aletaha D, Nell VP, 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-R806.
- Prevoost ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. 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.
- Disease Activity Score in rheumatoid arthritis. Nijmegen, the Netherlands: University Medical Center, 2016 (<http://www.das-score.nl/das28/>).
- Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med* 2009;28:586-604.
- Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med* 2014;33:693-713.

19. Jones G, Sebba A, Gu J, et al. Comparison of tocilizumab monotherapy versus methotrexate monotherapy in patients with moderate to severe rheumatoid arthritis: the AMBITION study. *Ann Rheum Dis* 2010;69:88-96.
20. Weinblatt ME, Schiff M, Valente R, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. *Arthritis Rheum* 2013;65:28-38.
21. Breedveld FC, Weisman MH, Kavanaugh AF, et al. The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54:26-37.
22. Curtis JR, Singh JA. Use of biologics in rheumatoid arthritis: current and emerging paradigms of care. *Clin Ther* 2011;33:679-707.
23. Moreland LW, O'Dell JR, Paulus HE, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of Early Aggressive Rheumatoid Arthritis Trial. *Arthritis Rheum* 2012;64:2824-35.
24. Nam JL, Ramiro S, Gaujoux-Viala C, et al. Efficacy of biological disease-modifying antirheumatic drugs: a systematic literature review informing the 2013 update of the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2014;73:516-28.

Copyright © 2017 Massachusetts Medical Society.

AN NEJM APP FOR iPhone

The NEJM Image Challenge app brings a popular online feature to the smartphone. Optimized for viewing on the iPhone and iPod Touch, the Image Challenge app lets you test your diagnostic skills anytime, anywhere. The Image Challenge app randomly selects from 300 challenging clinical photos published in NEJM, with a new image added each week. View an image, choose your answer, get immediate feedback, and see how others answered. The Image Challenge app is available at the iTunes App Store.

