Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Verstovsek S, Mesa RA, Gotlib J, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. N Engl J Med 2012;366:799-807.

Supplementary Appendix

Supplement to: Verstovsek S, Mesa R, Gotlib J, et al. A Double-Blind, Placebo-Controlled Trial of Ruxolitinib for Myelofibrosis

Contents:

COMFORT-I Investigators

The following investigators contributed to the study (listed in alphabetical order by country): **Australia**—P. Cannell, Royal Perth Hospital, Perth, WA; J. V. Catalano, Frankston Hospital and Department of Clinical Haematology, Monash University, Frankston, Melbourne, Victoria; B. H. Chong, St. George Hospital, Kogarah, NSW; P. Coughlin, Monash University/Box Hill Hospital, Box Hill, Victoria; S. T. S. Durrant, Royal Brisbane and Women's Hospital, Herston, Queensland; T. E. Gan, Monash Medical Centre, Clayton, Victoria; H. C. Lai, Townsville Hospital, Douglas, QLD; M. F. Leahy, Fremantle Hospital and Health Service, Fremantle, WA; M. Leyden, Maroondah Hospital, Ringwood East, Victoria; R. Lindeman, Prince of Wales Hospital, Randwick, NSW; D. Ma, St. Vincent's Hospital, Darlinghurst, NSW; A. Perkins, Haematology and Oncology Clinics of Australia, Milton, QLD; A. C. Perkins, Princess Alexandra Hospital, Woolloongabba, QLD; D. Ross, Flinders Medical Centre, Bedford Park, SA; W. Stevenson, Royal North Shore Hospital, St. Leonards, NSW; **Canada**—K. Grewal, Eastern Health, St. John's, NL; V. Gupta, Princess Margaret Hospital, University of Toronto, Toronto, ON; K. Howson-Jan, London Health Sciences Centre, London, ON; S. Jackson, St. Paul's Hospital, Vancouver, BC; C. Shustik, Royal Victoria Hospital, Montreal, QC; R. van der Jagt, Ottawa Hospital-General Campus, Ottawa, ON; **United States**—L. Afrin, Hollings Cancer Center, Charleston, SC; L. P. Akard, Indiana Blood and Marrow Transplantation, LLC, Beech Grove, IN; M. O. Arcasoy, Duke University Medical Center, Durham, NC; E. Atallah, Froedtert Hospital and Medical College of Wisconsin, Milwaukee, WI; J. Altman, Northwestern Memorial Hospital, Chicago, IL; J. Camoriano, Mayo Clinic Arizona, Scottsdale, AZ; T. P. Cescon, Berks Hematology Oncology Associates, West Reading, PA; C. R. Cogle, University of Florida, Gainesville, FL; R. Collins, Jr., University of Texas Southwestern Medical Center, Dallas, TX; K-

H. Dao, Oregon Health and Science University, Portland, OR; H. J. Deeg, Fred Hutchinson Cancer Research Center, Seattle, WA; M. Deininger, Oregon Health and Science University, Portland, OR; N. J. DiBella, Rocky Mountain Cancer Centers, Aurora, CO; J. F. DiPersio, Washington University School of Medicine, St. Louis, MO; A. Faitlowicz, University of California-Irvine Medical Center, Orange, CA; F. A. Fakih, Florida Pulmonary Research Institute, LLC, Winter Park, FL; R. Frank, Norwalk Hospital, Norwalk, CT; N. Y. Gabrail, Gabrail Cancer Center Research, Canton, OH; S. L. Goldberg, Hackensack University Medical Center, Hackensack, NJ; J. Gotlib, Stanford Cancer Institute, Stanford, CA; H. M. Gross, Dayton Physicians, LLC, Dayton, OH; J. H. Harvey, Jr., Birmingham Hematology and Oncology Associates, LLC, Birmingham AL; R. H. Herzig, University of Louisville, Louisville, KY; E. Hexner, Abramson Cancer Center at the University of Pennsylvania, Philadelphia, PA; C. E. Holmes, Vermont Cancer Center, Burlington, VT; E. Ibrahim, Beaver Medical Group, Highland, CA; R. Jacobson, Palm Beach Cancer Institute, West Palm Beach, FL; C. Jamieson, Moores University of California-San Diego Cancer Center, La Jolla, CA; K. Jamieson, University of Iowa Hospitals and Clinic, Iowa City, IA; C. M. Jones, Jones Clinic, PC, Germantown, TN; H. M. Kantarjian, University of Texas M.D. Anderson Cancer Center, Houston, TX; A. Kassim, Vanderbilt Clinic, Nashville, TN; C. M. Kessler, Georgetown University Medical Center, Washington, DC; T. Kindwall-Keller, University Hospitals Case Medical Center, Cleveland, OH; P. P. N. Lee, Tower Cancer Research Foundation, Beverly Hills, CA; R. M. Lyons, Cancer Care Centers of South Texas/US Oncology, San Antonio, TX; R. Marschke, Jr., Front Range Cancer Specialists, Fort Collins, CO; J. Mascarenhas, Mount Sinai School of Medicine, New York, NY; E. Meiri, Palm Beach Institute of Hematology and Oncology, Boynton Beach, FL; A. Menter, Kaiser Permanente, Denver, CO; R. A. Mesa, Mayo Clinic-Arizona, Scottsdale, AZ; C. Miller, St. Agnes HealthCare, Inc, Baltimore, MD; C. O'Connell, University of Southern California, Los Angeles, CA; I. Okazaki, Straub Clinic and Hospital, Honolulu, HI; R. Orlowski, Carolina Oncology Specialists, PA, Hickory, NC; R. Paquette, University of California-Los Angeles Medical

Hematology and Oncology, Los Angeles, CA; V. R. Phooshkooru, Mid Dakota Clinic, PC, Bismarck, ND; B. Powell, Wake Forest University Health Services, Winston-Salem, NC; J. T. Prchal, Huntsman Cancer Institute, Salt Lake City, UT; R. Ramchandren, Karmanos Cancer Institute, Detroit, MI; F. Rana, Shands Jacksonville Clinical Center, Jacksonville, FL; A. Raza, Columbia University Medical Center, New York, NY; C. Rivera, Mayo Clinic-Jacksonville, Jacksonville, FL; E. A. Sahovic, Western Pennsylvania Hospital, Pittsburgh, PA; M. Scola, Carol G. Simon Cancer Center, Morristown, NJ; M. Scouros, Houston Cancer Institute, PA, Houston, TX; M. Sekeres, Cleveland Clinic, Cleveland, OH; J. Shammo, Rush University Medical Center, Chicago, IL; R. S. Siegel, George Washington University, Washington, DC; R. T. Silver, Weill Cornell Medical Center, New York, NY; C. P. Spears, Sierra Hematology and Oncology, Sacramento, CA; M. Talpaz, University of Michigan Medical Center, Ann Arbor, MI; M. Tsai, Park Nicollet Institute, St. Louis Park, MN; S. Verstovsek, University of Texas M.D. Anderson Cancer Center, Houston, TX; T. Walters, Mountain States Tumor Institute, Boise, ID; R. S. Weiner, Arena Oncology Associates, PC, Lake Success, NY; E. F. Winton, Emory University Hospital, Atlanta, GA; S. E. Young, Somerset Hematology-Oncology Associates, Somerville, NJ; F. Yunus, University of Tennessee Cancer Institute, Memphis, TN.

Study Methods

Inclusion and Exclusion Criteria

Patients ≥18 years of age with primary myelofibrosis, post–polycythemia vera myelofibrosis, or post–essential thrombocythemia myelofibrosis (according to World Health Organization criteria revised in 2008)[1]; a life expectancy of at least 6 months; an International Prognostic Scoring System (IPSS; Supplementary Table S1) score of 2 (intermediate-2 risk) or ≥3 (high risk); and an Eastern Cooperative Oncology Group (ECOG) performance status of ≤3 were enrolled. The ECOG performance status is on a scale of 0 to 5, with higher scores indicating greater disability; 5=dead and 3=a capacity for only limited self-care and confinement to a bed or chair for more than 50% of waking hours (http://www.ecog.org/general/perf_stat.html). The IPSS assigns a value of 1 for each of the following prognostic factors, if present: age >65 years, presence of constitutional symptoms (weight loss, fever, night sweats), anemia (hemoglobin <100 g/l), leukocytosis (history of white blood cell count >25x10^9/l), and circulating blasts ≥1%.[2] Patients were also required to have a peripheral blast count of <10%, an absolute peripheral blood CD34+ cell count >20x10^6/l, and palpable splenomegaly of ≥5 cm below the left costal margin. Patients had to have disease that was resistant or refractory to available treatment or to be intolerant of or not candidates for such therapy, and to have disease that required treatment, as defined by any 1 of the following: IPSS prognostic score of ≥3 (high risk), palpable splenomegaly of at least 10 cm below the left costal margin, or a score of at least 3 (on a scale of 0 [absent] to 10 [worst imaginable]) on at least 2 items or a score of 5 on 1 item of a screening version of the modified Myelofibrosis Symptom Assessment Form (MFSAF) v2.0 diary.

Supplementary Table S1. IPSS Scoring System.

IPSS Risk Factors*
Age >65 years
White cell count >25 x 10^9/l
Hemoglobin <10 g/d
Peripheral blood blasts ≥1%
Constitutional symptoms

*The presence of each factor is assigned 1 point:

- Low risk: 0 points
- Intermediate-1 risk: 1 point
- Intermediate-2 risk: 2 points
- High risk: 3 or more points

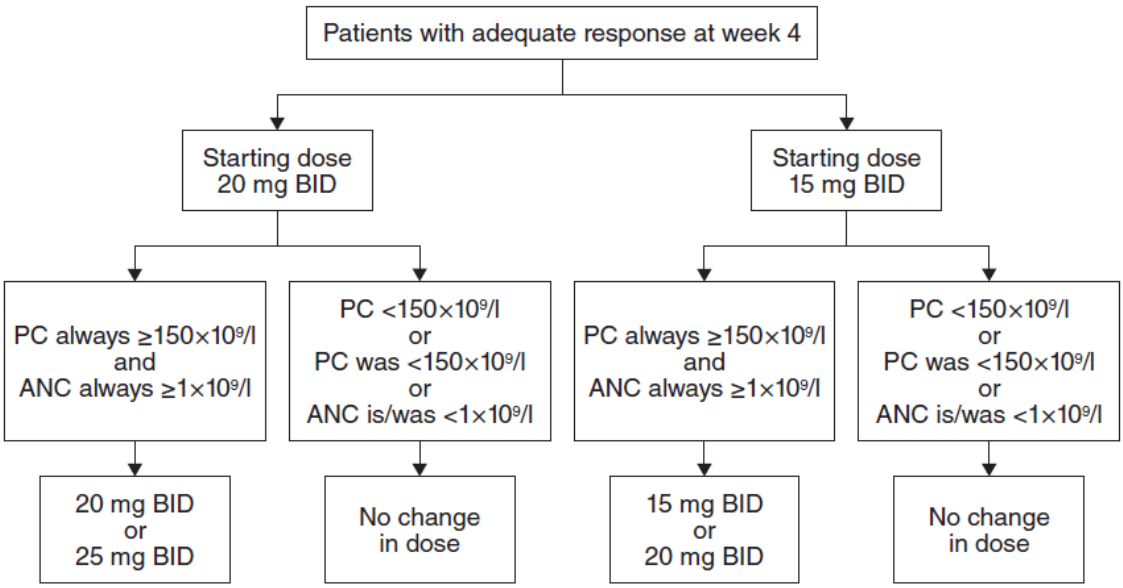
Patients were excluded from the study for an absolute neutrophil count (ANC) ≤1x10^9/l, platelet count <100x10^9/l, direct bilirubin ≥2x the laboratory upper limit of normal (ULN), alanine aminotransferase ≥2.5x the laboratory ULN, creatinine >2.0 mg/dl, history of malignancy within the previous 5 years (apart from cured basal cell or squamous cell carcinoma of the skin), splenic irradiation within 12 months prior to randomization, and any prior therapy with JAK inhibitors. Other investigational agents must have been discontinued 14 days or 6 half-lives prior to the initial baseline visit, and any other treatment for myelofibrosis, including hydroxyurea, interferon, thalidomide, busulfan, lenalidomide, and anagrelide, must have been discontinued within 28 days prior to the first baseline visit.

Dose Adjustments

The initial dose of ruxolitinib was determined by the baseline platelet count. After the first 4 weeks of therapy, dose regimens could be increased by 5 mg twice daily in patients who demonstrated inadequate efficacy and who met the following 3 criteria (Supplementary Figure S1):

1. Inadequate efficacy demonstrated by palpable spleen length below the left costal margin that had been reduced by <40% at the week 4 visit relative to baseline
2. Platelet count at week 4 was ≥150x10^9/l and platelet count was never <150x10^9/l at a prior laboratory evaluation since baseline
3. ANC levels remained at ≥1x10^9/l since baseline

Supplementary Figure S1. Rules for Dose Increases in Ruxolitinib or Placebo because of Inadequate Efficacy.



ANC denotes absolute neutrophil count, BID twice daily, PC platelet count.

Administration of active study medication or placebo was to be interrupted if platelet counts declined below 50x10^9/l (grade 3 laboratory abnormality by Common Terminology Criteria

for Adverse Events v3.0) or if the ANC fell below $0.5x10^9$/l. Doses were decreased for platelet counts <$125x10^9$/l, as shown in Supplementary Table S2. In order to provide sufficient data to make dose adjustment decisions, it was recommended that hematology parameters be obtained at least weekly for platelet counts <$100x10^9$/l or an ANC <$1x10^9$/l, and at least twice weekly for a platelet count <$50x10^9$/l or an ANC <0.5 $x10^9$/l.

Supplementary Table S2. Mandatory Dose Reductions in Ruxolitinib or Placebo Groups Based on Platelet Counts.

Platelet Count at Time of Decline	Dose at the Time of Platelet Decline				
	25 mg BID	20 mg BID	15 mg BID	10 mg BID	5 mg BID
	Dose That MUST be Instituted				
≥$125x10^9$/l	No dose reduction required				
100 to <125 $x10^9$/l	20 mg BID	20 mg BID	15 mg BID	10 mg BID	5 mg BID
75 to <100 $x10^9$/l	10 mg BID	10 mg BID	10 mg BID	10 mg BID	5 mg BID
50 to <75 $x10^9$/l	5 mg BID	5 mg BID	5 mg BID	5 mg BID	5 mg BID
<50 $x10^9$/l	Must hold administration				

BID denotes twice daily.

Dosing could be restarted or increased following recovery of platelet counts to acceptable levels. The objective for restarting or escalating after a reduction for safety was to find the highest safe dose of ruxolitinib for each patient, with increases generally in increments of no more than 5 mg twice daily and not more often than every 2 weeks. Those restarting after a reduction for thrombocytopenia could not receive doses higher than 20 mg twice daily. ANC levels that declined to <$0.5x10^9$/l necessitated immediate dose interruption. Recovery of ANC levels to >$0.5x10^9$/l but <$0.75x10^9$/l allowed restarting ruxolitinib at 5 mg twice daily; those with ANC levels between $0.75x10^9$/l and <$1x10^9$/l could restart treatment at 10 mg twice daily. Those with ANC increases to ≥$1x10^9$/l that were maintained could receive dose increases to a maximum of 20 mg twice daily (Supplementary Table S3).

Supplementary Table S3. Restarting or Increasing Ruxolitinib or Placebo after Safety Interruptions or Dose Reductions.

Current Platelet Count	Dose Restart or Dose Increase Guidelines (maximum doses)
<50 x10^9/l	Continue hold
50 to <75 x10^9/l	5 mg BID for 2 wk; if stable, may increase to 10 mg BID
75 to <100 x10^9/l	10 mg BID for 2 wk; if stable, may increase to 15 mg BID
100 to <125 x10^9/l	15 mg BID
≥125 x10^9/l	20 mg BID
Current ANC Level	**Dose Restart or Dose Increase Guidelines (maximum doses)**
<0.5 x10^9/l	Continue hold
0.5 to <0.75 x10^9/l	5 mg BID for 2 wk; if stable, may increase to 10 mg BID
0.75 to <1 x10^9/l	10 mg BID for 2 wk; if stable, may increase to 15 mg BID
>1x10^9/l	15 mg BID
>1.5 x10^9/l	20 mg BID

ANC denotes absolute neutrophil count, and BID twice daily.

Crossover Criteria

Patients with a ≥25% increase in spleen volume from baseline were eligible for early unblinding, and those receiving placebo were eligible for crossover to ruxolitinib treatment prior to the prespecified unblinding and analysis of the trial. Prior to week 24, such patients also had to demonstrate worsening early satiety accompanied by weight loss or worsening splenic pain demonstrated by increased narcotic requirements. Patients on ruxolitinib could remain on ruxolitinib therapy if spleen growth was asymptomatic and occurred after week 24. Patients obtaining benefit from therapy could continue to receive ruxolitinib until the later of marketing approval or when the last randomized patient completed week 144, provided they did not meet a criterion for withdrawal from the study.

Assessments

Spleen Volume: Spleen volume was assessed by magnetic resonance imaging (MRI), or by computed tomography for patients in whom MRI was contraindicated or in facilities where MRI was not readily available. The same modality was used for each patient throughout the study. Scans from individual patients were read by a central reviewer who was blinded to the patient's treatment assignment and clinical data. Spleen volume was calculated using Alice™ image analysis software at the central review contractor, Perceptive Infomatics®. Imaging for spleen volume assessment was obtained at baseline; at weeks 12, 24, 36, 48, 60, and 72; and every 24 weeks thereafter. Investigators were not provided with the results of the volume assessments.

Symptoms: Myelofibrosis symptoms were assessed using the modified MFSAF v2.0 diary. This diary was developed in accordance with the United States Food and Drug Administration (FDA) guidance on patient-reported outcome measures and in consultation with the Study Endpoints and Labeling Development Division of the FDA, and is a modified version of the MFSAF designed and validated by Mesa and colleagues.[3,4] The diary was provided to patients on an electronic handheld device on which they recorded their answers to 7 questions about their myelofibrosis symptoms. The questions assessed night sweats, itching (pruritus), abdominal discomfort, pain under the ribs on the left side, feeling of fullness (early satiety), muscle/bone pain, and inactivity on a scale of 0 (absent) to 10 (worst imaginable). The Total Symptom Score (TSS) was the sum of the individual symptom scores, excluding the score for inactivity, which was analyzed separately. Patients completed the diary daily for 7 days prior to starting the study drug to obtain a mean baseline value, and then daily through week 24. The baseline TSS was the mean of the daily scores through the 7-day baseline period. The 24-week TSS was the mean of the scores obtained during the 28 days prior to the week 24 visit.

Other Patient-Reported Outcome Measures: Other patient-reported outcome assessments performed at each visit included a Patient Global Assessment of Change (PGIC), which assesses a patient's overall impression of change on a scale of 1 (very much improved) to 7 (very much worse), with 4 representing no change. Other validated instruments included the European Organisation for Research and Treatment of Cancer Quality of Life 30 Questionnaire (EORTC QLQ-C30) and the Patient-Reported Outcomes Measurement System (PROMIS) Fatigue Scale. The EORTC QLQ-C30 is a common quality-of-life measure used for patients with cancer in clinical studies. It is a self-administered questionnaire that has 5 functional subscales (physical, role, emotional, cognitive, and social) and individual symptom scales (eg, fatigue, pain, and nausea). The PROMIS Fatigue Scale contains 7 items that measure the frequency or the impact of fatigue, including impact on daily activities. Each of the items uses a 5-point response option with scores of 1 (never) to 5 (always).

Transfusion Dependence and Independence: Transfusion status was assessed using 2 methods. With the protocol-specified method, transfusion dependence at baseline was defined as requiring ≥2 units of red blood cell (RBC) products over an 8-week period prior to the screening visit date. New-onset transfusion dependency was defined as the use of ≥2 units of RBC products during the final 8 weeks of a patient's participation prior to the data-cutoff date of the study in a patient who was not transfusion-dependent at baseline. Similarly, new-onset transfusion independence was defined as the requirement of 0 units of RBC products during the final 8 or 12 weeks of a patient's participation prior to the data-cutoff date in a patient who was transfusion-dependent at baseline. In addition, new-onset transfusion independence in initially dependent patients was assessed using criteria published by the International Working Group for Myelofibrosis Research and Treatment.[2] Baseline transfusion dependence was defined as the receipt of ≥2 units of RBC products in the 4 weeks prior to randomization, and on-study

transfusion independence was defined as the absence of transfusions for any period of ≥8 weeks.

Terms of Bleeding and Bruising: In the evaluation of safety, bleeding events included the following terms: blood urine present, conjunctival hemorrhage, epistaxis, gastric varices hemorrhage, gastrointestinal hemorrhage, genital hemorrhage, gingival bleeding, hematochezia, hematuria, hemoptysis, hemorrhage, hemorrhoidal hemorrhage, melena, post procedural hemorrhage, retinal hemorrhage, retroperitoneal hematoma, splenic hematoma, splenic hemorrhage, subdural hematoma. Bruising was explored separately and included the following terms: contusion, ecchymosis, hematoma, increased tendency to bruise, injection site hematoma, vessel puncture site hematoma, purpura, petechiae, and periorbital hematoma.

Statistical Analysis

In order to provide sufficient safety data and adequate power for secondary efficacy endpoints evaluating symptoms, the study was originally designed to enroll a total of 240 patients. However, eligible patients in active screening at the time that enrollment was suspended were allowed to enter the study, resulting in more patients than the 240 planned. Based on the assumption that at least 30% of patients in the ruxolitinib group and no more than 10% in the placebo group would achieve a ≥35% reduction in spleen volume at 24 weeks, 240 patients would provide 97% power to detect a treatment difference in spleen volume response at a 2-sided alpha level of 0.05 using a chi-square test.

The primary endpoint was analyzed using the Fisher exact test, as there were fewer than 4 responders in the placebo group. Comparative secondary efficacy variables were tested in a fixed-sequence-testing procedure at an alpha level of 0.05 in the following order: proportion of patients with ≥50% reduction in TSS from baseline to week 24 (using chi-square test), actual

change in TSS from baseline to week 24 (using both Wilcoxon signed rank-sum test and analysis of covariance), and overall survival (using log-rank test). A post hoc analysis comparing baseline characteristics between groups was conducted using the Wilcoxon rank-sum test for continuous variables and chi-square test for categorical variables.

Analyses were conducted in accordance with intent-to-treat (ITT) principles. For all applicable variables, however, patients with missing baseline values were excluded from comparisons of baseline characteristics between groups and the analysis of change and percent change from baseline. Patients who discontinued prior to week 24 or crossed over prior to week 24 were counted as nonresponders (for response measures of spleen volume reduction and symptom improvement from baseline to week 24).

Durability of spleen response was analyzed using the Kaplan-Meier method. This analysis included all patients who had at least 1 assessment of spleen volume demonstrating a ≥35% reduction from baseline. The start date was the date of the first imaging demonstrating a ≥35% reduction in spleen volume from baseline. To address an oversight in the original SAP that a trivial increase in spleen volume from at least a 35% reduction to less than a 35% reduction could be classified as a loss of response, the method used for determining durability of spleen response reported here was defined prior to review of any spleen volume data but after the protocol and SAP were finalized. This method, which is consistent in approach to other definitions of response duration in oncology, differs from the SAP-defined method primarily by defining loss of response as a reduction in spleen volume <35% from baseline that is also a 25% increase over nadir. Duration of response = (end date – start date) + 1 day. For patients with no end date, duration was censored at the date of the last adequate assessment of the spleen volume. The analysis was conducted at the time of data cutoff.

Interaction analyses were performed using the ANCOVA method, controlling for baseline and subgroups (specified in the statistical analysis plan). For the percentage change from baseline to week 24 in spleen volume, the baseline spleen volume and palpation length were used as the covariates; for percentage change from baseline in TSS, the baseline TSS was used as the covariate.

Survival time was analyzed using the Kaplan-Meier method, according to the original randomization group, regardless of treatment crossover for all ITT patients. Two analyses were conducted. The first analysis was at the time of the prospectively defined primary-analysis cutoff date for the study (November 2, 2010), and the second was performed at the time of the data cutoff for a prospectively defined 4-month safety follow-up (March 1, 2011), required as part of a New Drug Application to the FDA.[5] For both analyses, the date of first study dose was used as the origin (beginning) for the survival-time calculation. For patients who were still on study at the data cutoff, the survival time was censored at the cutoff date. For patients who discontinued from the study, sites were requested to obtain patient survival information every 6 months, and the survival time was censored at the last available contact date if the patient was still alive. For patients who died either on study or after discontinuation based on the follow-up information, the death was considered as an uncensored event, and the survival time was the difference between the origin and the death date plus 1 day.

Adverse events and serious adverse events were reported using National Cancer Institute Common Terminology Criteria for Adverse Events version 3.0 and were tabulated along with other safety data, including laboratory and electrocardiographic data. Adverse events of new onset or increased grade during dose interruptions were also summarized.

References

1. Tefferi A, Vardiman JW. Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. Leukemia 2008;22:14-22.

2. Tefferi A, Barosi G, Mesa RA, et al. International Working Group (IWG) consensus criteria for treatment response in myelofibrosis with myeloid metaplasia, for the IWG for Myelofibrosis Research and Treatment (IWG-MRT). Blood 2006;108:1497-503.

3. Mesa RA, Kantarjian H, Tefferi A, et al. Evaluating the serial use of the Myelofibrosis Symptom Assessment Form for measuring symptomatic improvement: performance in 87 myelofibrosis patients on a JAK1 and JAK2 inhibitor (INCB018424) clinical trial. Cancer 2011;117:4869-77.

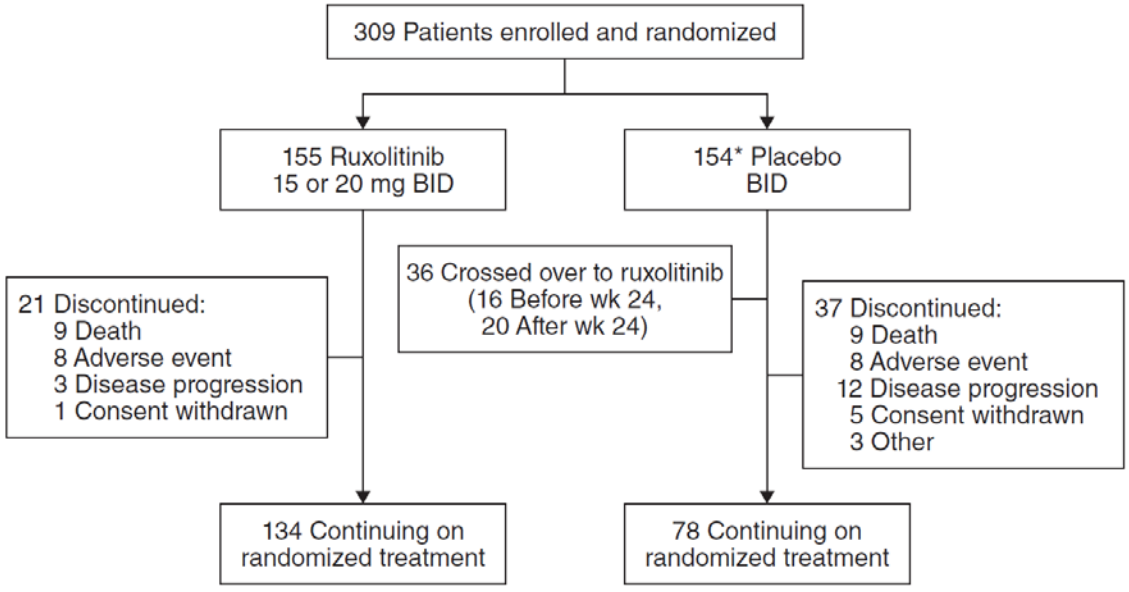
4. Mesa RA, Kantarjian H, Tefferi A, et al. Validation of the serial use of the Myelofibrosis Symptom Assessment Form (MF-SAF) for measuring symptomatic improvement: performance in 86 myelofibrosis patients on INCB018424 clinical trial. Blood 2009;114:3917. Abstract.

5. US Food and Drug Administration. Attachment B: Clinical Safety Review of an NDA or BLA of the Good Review Practice: Clinical Review Template (MAPP 6010.3 Rev. 1).

Results Not Included in Main Paper

Supplementary Figure S2. Patient Disposition.

Figure S2 shows patient disposition over the time course of the study, including primary reasons for treatment discontinuation. The data-cutoff date occurred when half of the patients remaining in the study (including those continuing on randomized treatment and those who had crossed over to ruxolitinib from placebo) completed the week-36 visit, and all patients enrolled had completed week 24 or discontinued. A total of 212 patients (134 ruxolitinib and 78 placebo) remained in the study as of the November 2, 2010 data-cutoff date. Of the patients who did not cross over to ruxolitinib, more withdrew from the placebo group than from the ruxolitinib group. Primary reasons for withdrawal from the study for these patients were death, adverse events, and disease progression.



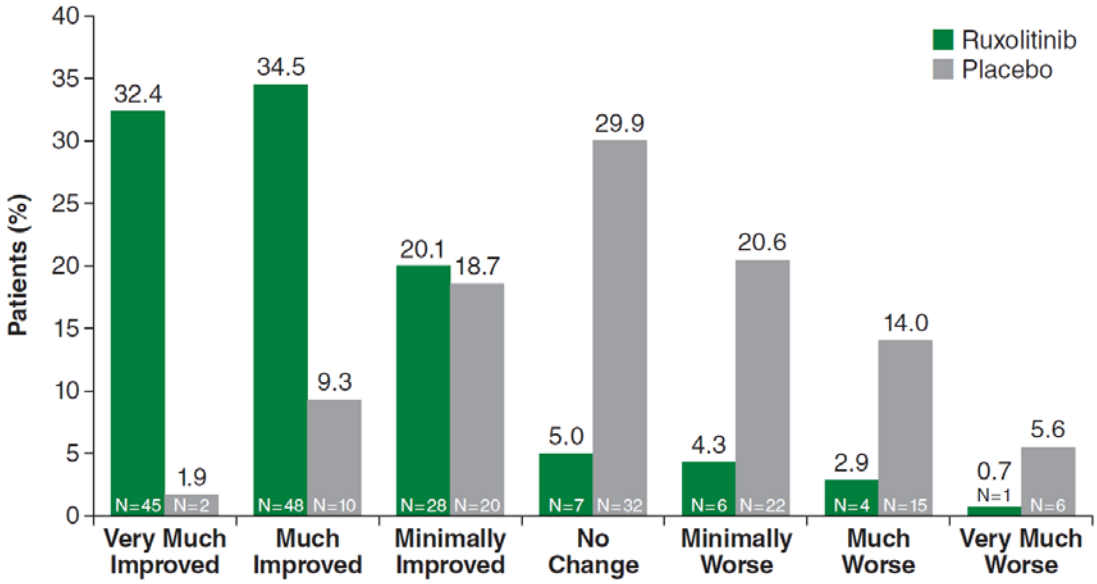
*3 patients not evaluable for safety but included in the intent-to-treat analysis of efficacy. BID denotes twice daily.

Twenty-four enrolled patients died by the time of data cutoff. Of these, 20 patients died during the study or within 28 days of the last dose of study medication, and 4 died more than 28 days after withdrawing from study medication. Of these 20 patients, 9 (5.8%) were randomized to ruxolitinib and 11 (7.3%) to placebo. Although 11 patients died in the placebo group, 9 deaths are noted in disposition of the placebo group for the following reasons: 1 of the 11 deaths occurred after the patient withdrew from the study but within the 28-day safety follow-up period, and 1 died after having crossed over to ruxolitinib.

An additional 4 months of data were collected after the cutoff date as part of a planned safety follow-up. At the time of this extended follow-up, there were a total of 13 deaths in the ruxolitinib group and 24 in the placebo group. At this time, all except 2 patients in the placebo group had crossed over to ruxolitinib.

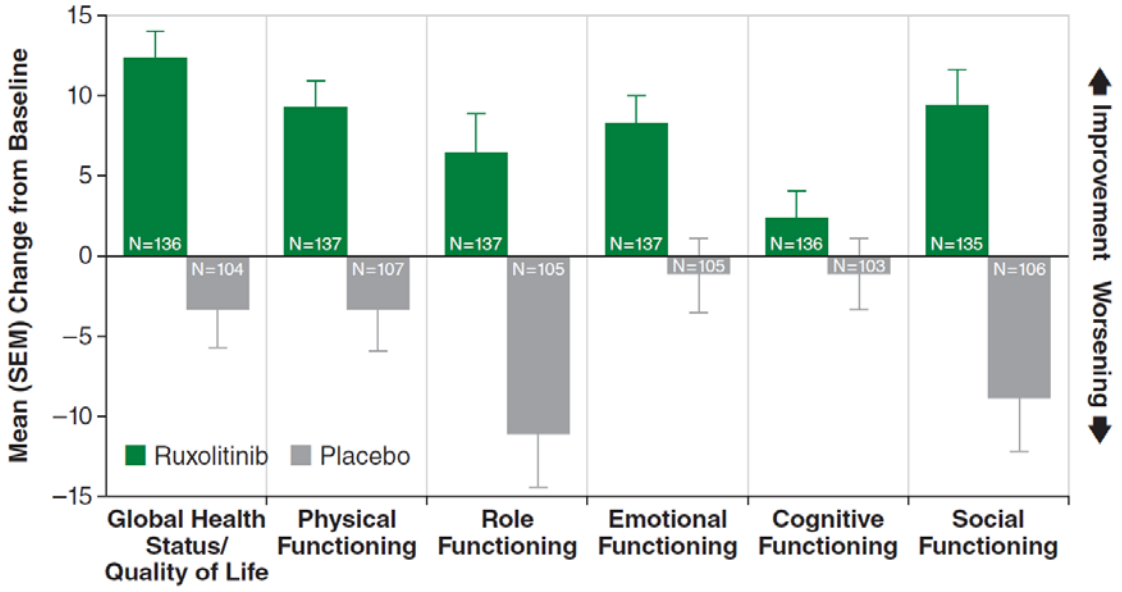
Supplementary Figure S3A. **Patient Global Impression of Change (PGIC) Scores at Week 24.**

At week 24, the majority of ruxolitinib-treated patients rated their condition as much improved or very much improved, whereas most placebo-treated patients rated their condition as unchanged or worse.



Supplementary Figure S3B. Mean Change from Baseline in EORTC QLQ-C30 Global Health Status and Functional Scales Results at Week 24.

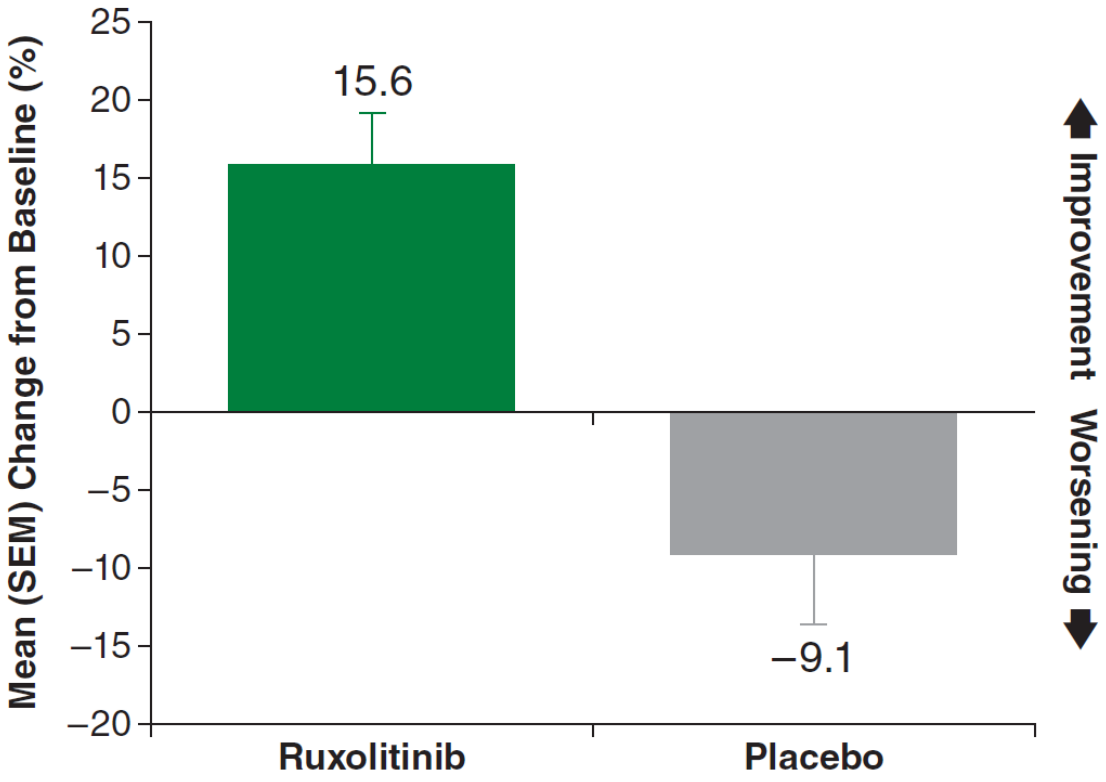
At week 24, the ruxolitinib group showed significant improvement in the global health status and functioning subscales of the European Organisation for Research and Treatment of Cancer Quality of Life 30 Questionnaire (EORTC QLQ-C30) compared with the placebo group, with a P value of <0.0001 reported for all subscales other than emotional functioning (P=0.0009) and cognitive functioning (P=0.06). For each subscale, ruxolitinib-treated patients exhibited improvement from baseline, whereas placebo-treated patients showed a worsening from baseline.



SEM denotes standard error of the mean.

Supplementary Figure S3C. Mean Percent Change in PROMIS Fatigue Scale at Week 24.

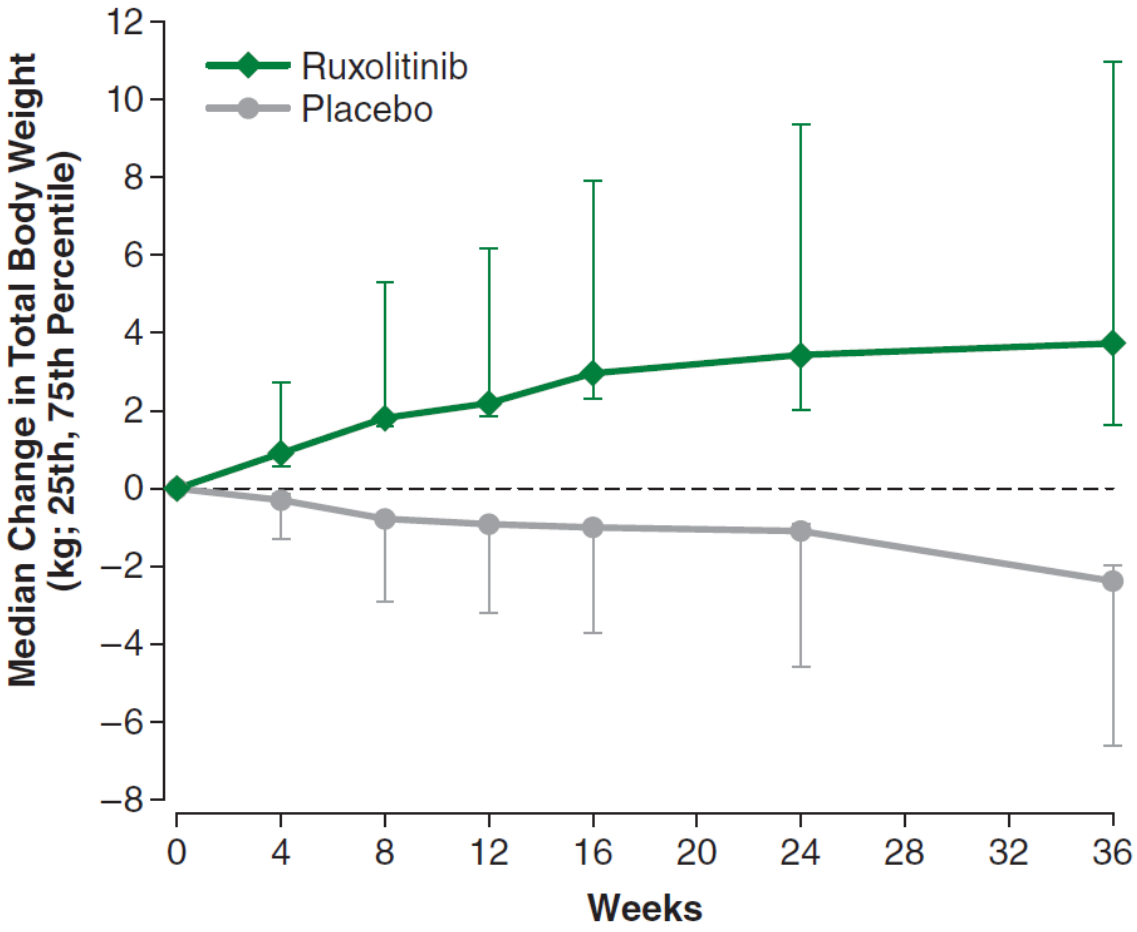
At week 24, the ruxolitinib group displayed a significantly greater mean percent improvement from baseline in fatigue compared with the placebo group (15.6% improvement in the ruxolitinib group and 9.1% worsening in the placebo group; P<0.0001).



PROMIS denotes Patient-Reported Outcomes Measurement System, and SEM standard error of the mean.

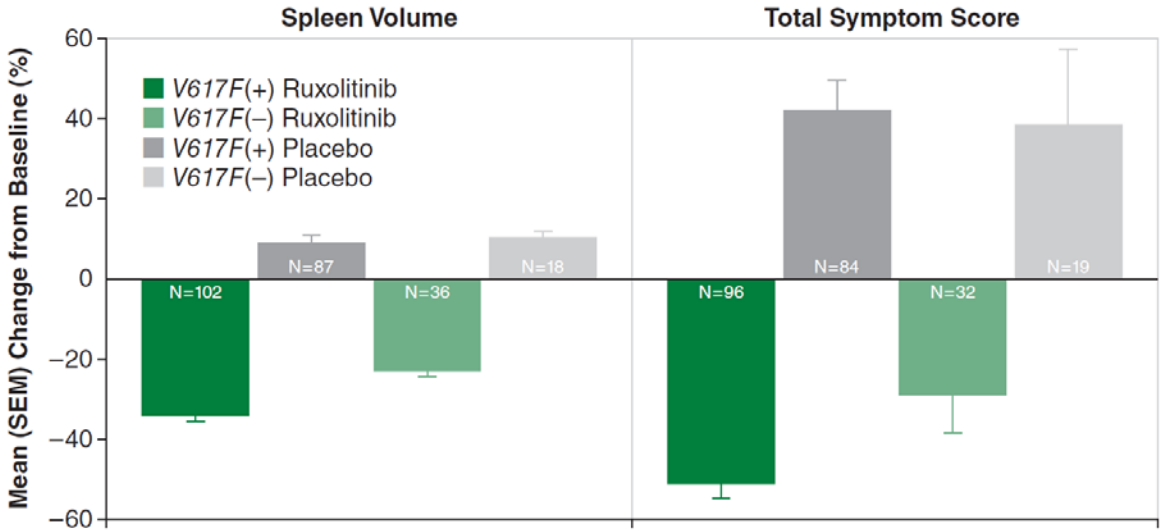
Supplementary Figure S4. Changes in Body Weight Over Time.

Consistent with the changes in spleen volume, symptoms, leptin, and inflammatory cytokines, ruxolitinib-treated patients experienced an increase in body weight over time, whereas those receiving placebo lost weight.



Supplementary Figure S5A. Changes from Baseline to Week 24 in Spleen Volume and Total Symptom Score in Patients With and Without the *JAK2V617F* Mutation.

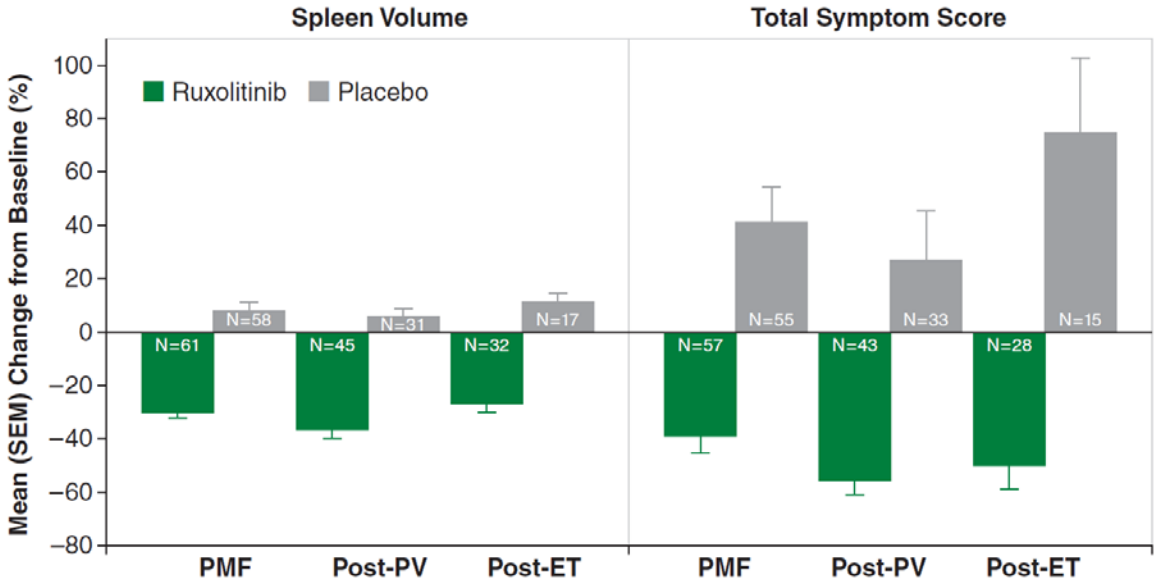
Patients treated with ruxolitinib with and without the *JAK2V617F* mutation achieved similar reductions in spleen volume compared with placebo. A similar trend was observed in Total Symptom Score on the Myelofibrosis Symptom Assessment Form. Therefore, ruxolitinib efficacy does not depend on the presence or absence of the *JAK2V617F* mutation.



SEM denotes standard error of the mean.

Supplementary Figure S5B. Changes from Baseline to Week 24 in Spleen Volume and Total Symptom Score in Patients With Primary Myelofibrosis, Post–Polycythemia Vera Myelofibrosis, and Post–Essential Thrombocythemia Myelofibrosis.

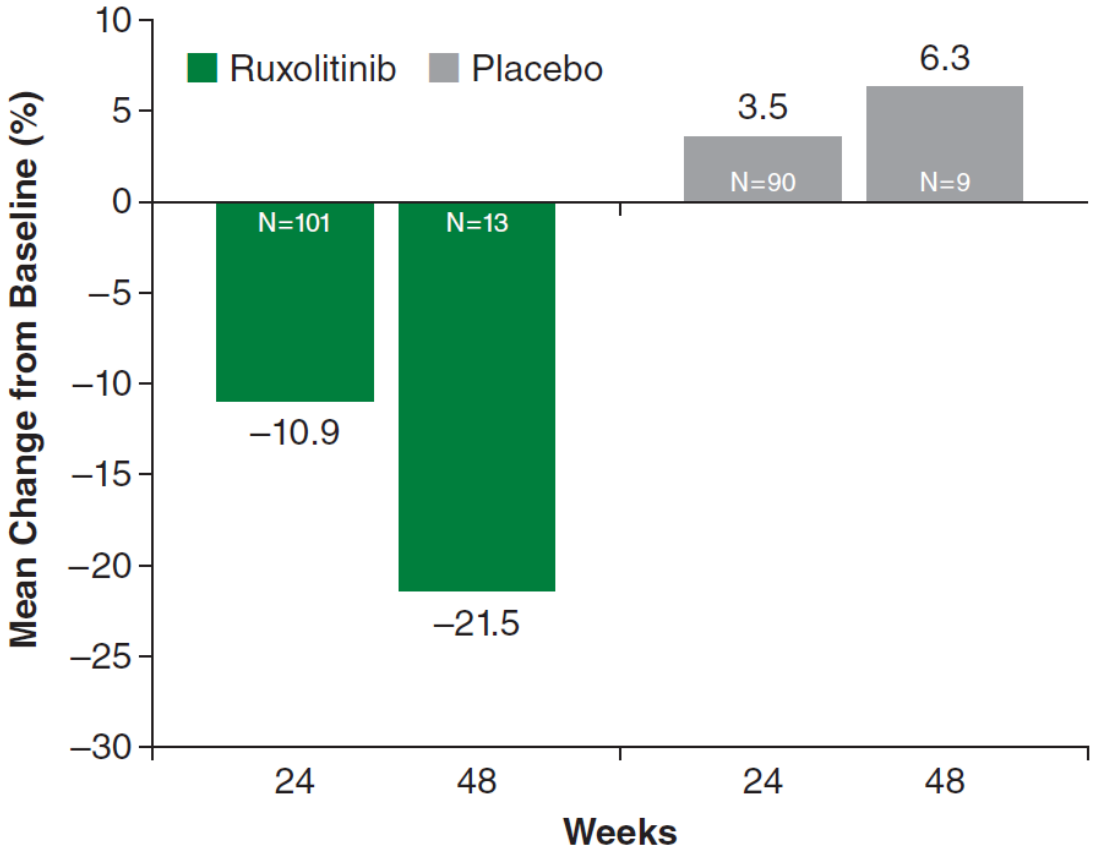
Patients treated with ruxolitinib achieved similar reductions in spleen volume and improvements in Total Symptom Score regardless of disease subtype.



ET denotes essential thrombocythemia (myelofibrosis), PMF primary myelofibrosis, PV polycythemia vera (myelofibrosis), and SEM standard error of the mean.

Supplementary Figure S6. Mean Percent Change from Baseline in Percent *JAK2V617F* at Weeks 24 and 48.

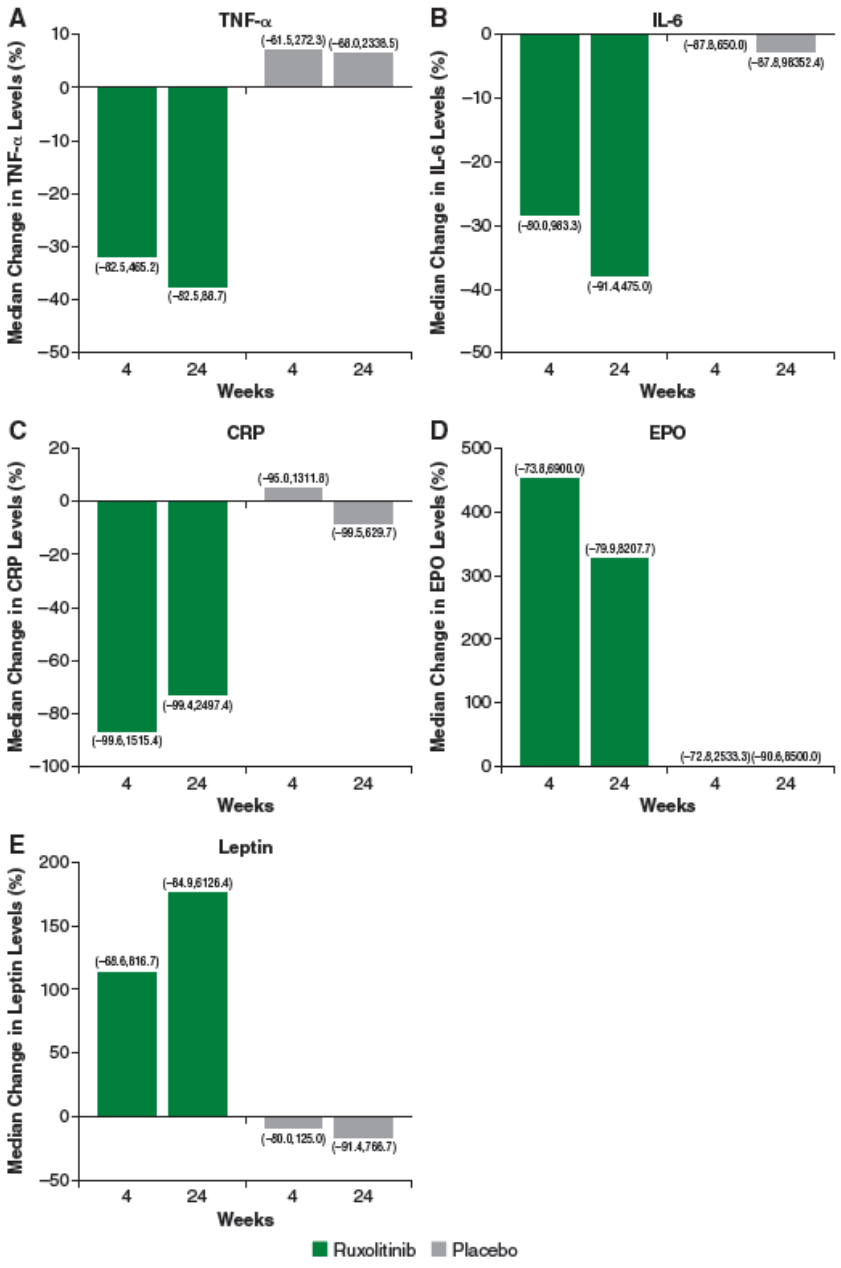
The percentage of *JAK2V617F* mutant allele relative to total *JAK2* (wild type *JAK2* plus *JAK2V617F*, referred to as *JAK2V617F* allele burden) was assessed. *JAK2V617F* allele burden was measured in whole blood at baseline and weeks 24 and 48 using a previously published method (Levine RL et al. Blood 2006;107:4139-41). Patients in the ruxolitinib group had a mean percent decrease in *JAK2V617F* at weeks 24 and 48, whereas those in the placebo group had mean percent increases at both time points. In the ruxolitinib group, the mean percent change from baseline was –10.9% at week 24 (N=101; P<0.0001 from the rank test) and –21.5% at week 48 (N=13; P=0.0002 from the rank test). The respective median values (ranges) were –7.8% (–83.7% to 35.7%) and –16.8% (–91.3% to –2.2%). This is in contrast to a mean percent change of 3.5% in the placebo group at week 24 (N=90; P=0.02 from the rank test) and a nonstatistically significant mean percent change from baseline of 6.3% at week 48 in the placebo group (N=9). The respective median values (ranges) were 1.1% (–23.3% to 100%) and –1.1% (–13.0% to 83.3%). Changes in allele burden were independent of myelofibrosis subtype and baseline *JAK2V617F* level.


Ruxolitinib
Placebo
Mean Change from Baseline (%)
10
5
0
−5
−10
−15
−20
−25
−30
N=101
N=13
−10.9
−21.5
3.5
6.3
N=90
N=9
24
48
24
48
Weeks

Supplementary Figure S7. Changes in Inflammatory Cytokines, Erythropoietin, and Leptin.*

Plasma markers associated with the symptoms of myelofibrosis were evaluated. Baseline elevations in inflammatory markers such as tumor necrosis factor-α (TNF-α; A), interleukin-6 (IL-6; B), and C-reactive protein (CRP; C) were noted in patients with myelofibrosis compared with healthy subjects. Decreases in these markers were observed over the 24 weeks of treatment with ruxolitinib, consistent with the observed changes in these symptoms. Erythropoietin (EPO; D) signals through JAK2, and levels of EPO increased following ruxolitinib treatment.

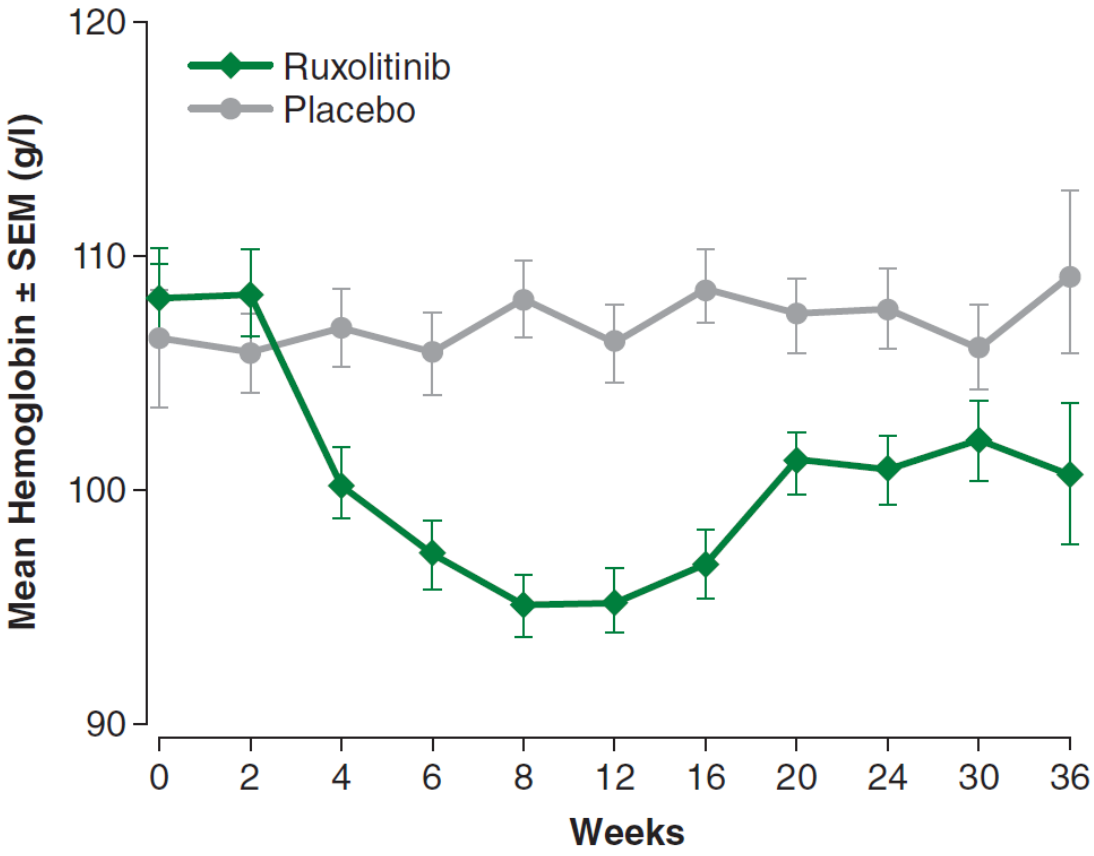
Plasma leptin (E), which signals through JAK2, has been shown to be a measure of the amount of body fat, and most individuals with myelofibrosis present with a state of catabolic excess. Treatment with ruxolitinib resulted in a greater than 2-fold increase in the median levels of plasma leptin at 4 weeks, and the increase continued over the 24 weeks. In contrast, leptin levels in the placebo group declined slightly over the 24 weeks.



*Ranges are shown above the bars in parentheses.

Supplementary Figure S8. Mean Hemoglobin Over Time.

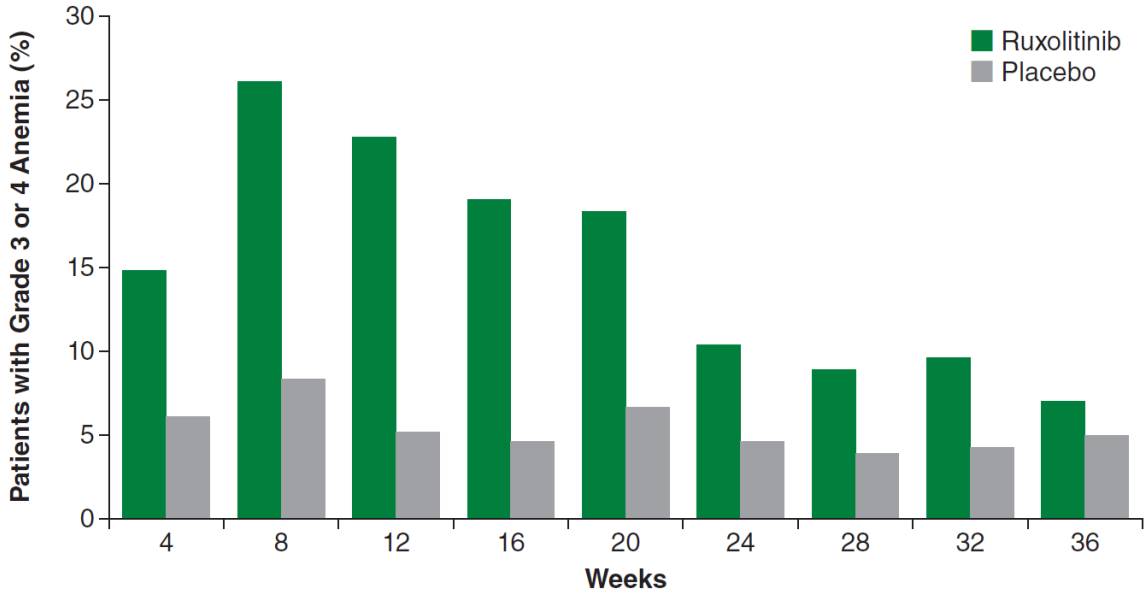
Mean hemoglobin levels over time in the ruxolitinib group reached a nadir of 95 g/l at approximately 8 to 12 weeks and then recovered to a new steady state by week 24. This pattern of a nadir followed by a recovery to a new steady state was seen in patients receiving or not receiving transfusion and independently of dose modification.



SEM denotes standard error of the mean.

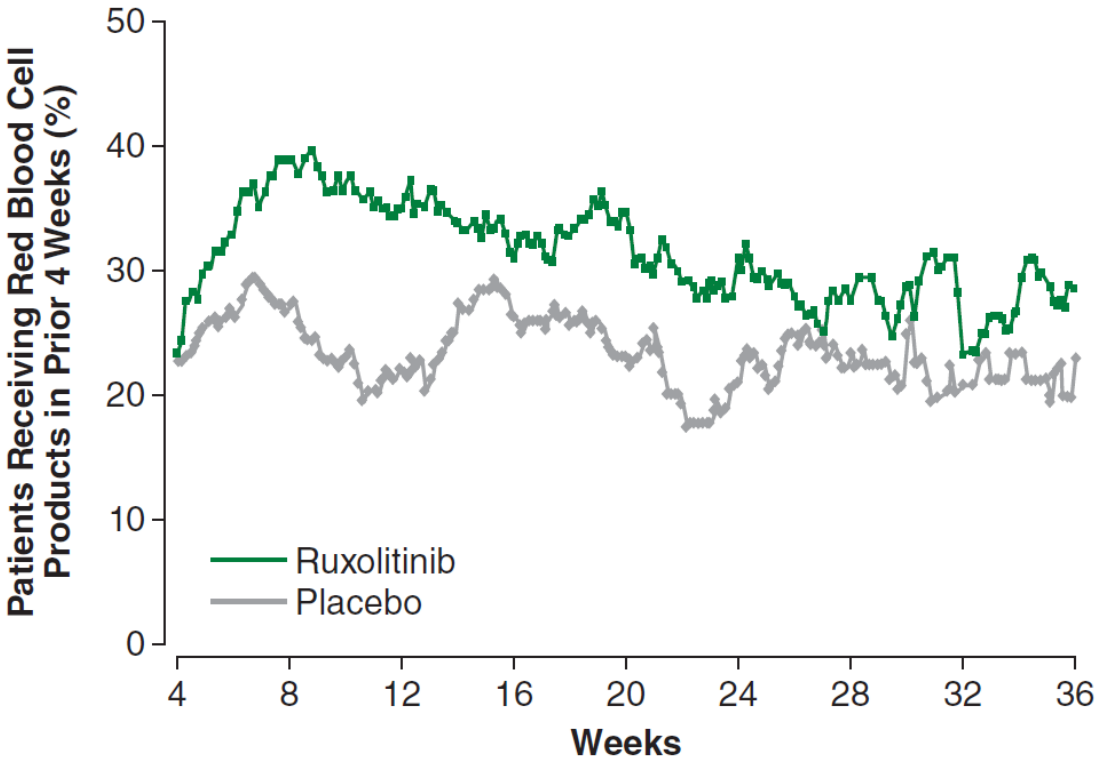
Supplementary Figure S9A. Monthly Prevalence of New-Onset or Continuing Grade 3 or 4 Anemia.

The monthly prevalence of grade 3 or 4 anemia increased over the first 8 to 12 weeks on ruxolitinib and gradually decreased over time to levels similar to placebo.



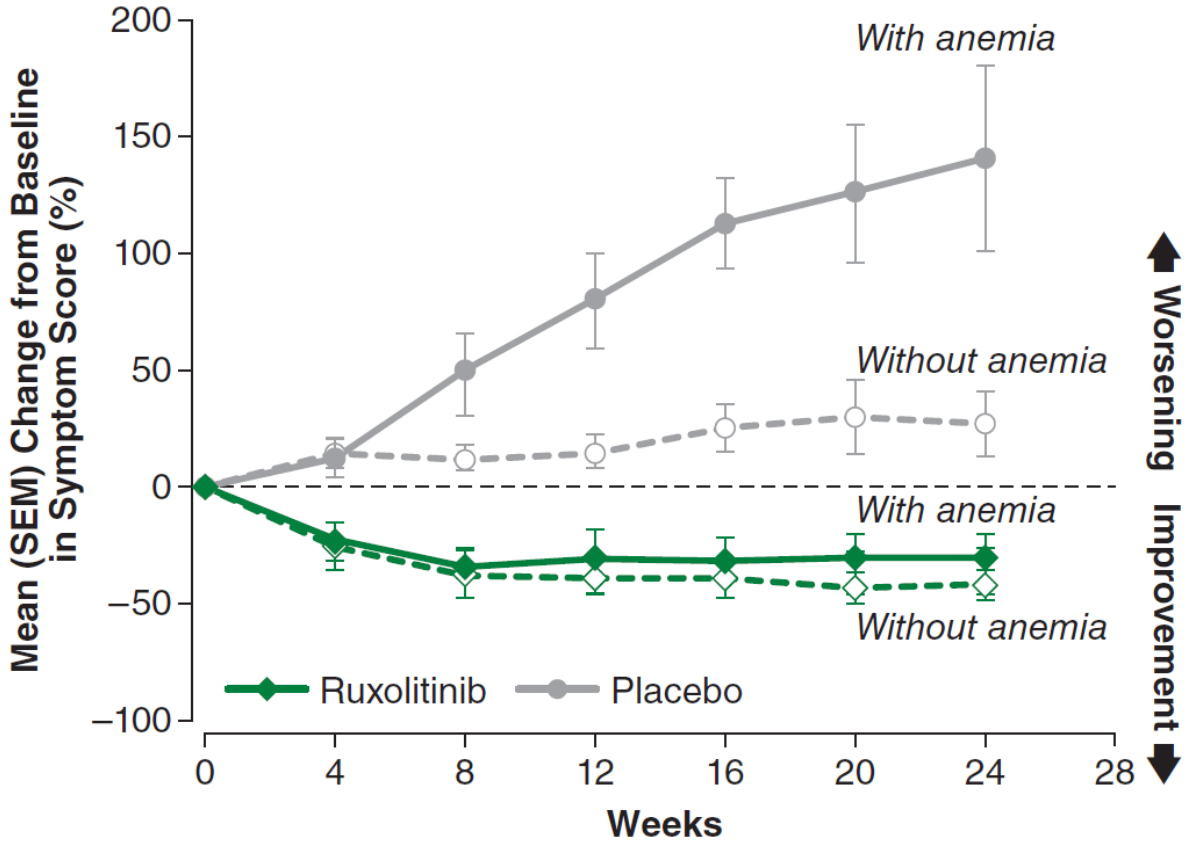
Supplementary Figure S9B. Patients Requiring Red Blood Cell Transfusions.

The proportion of patients requiring transfusions (1 or more units of red blood cells) increased over the first 8 to 12 weeks of ruxolitinib treatment, then decreased to levels similar to those with placebo by week 24.



Supplementary Figure S9C. Effect of Grade 3 or 4 Anemia on Symptom Score Changes Over Time.

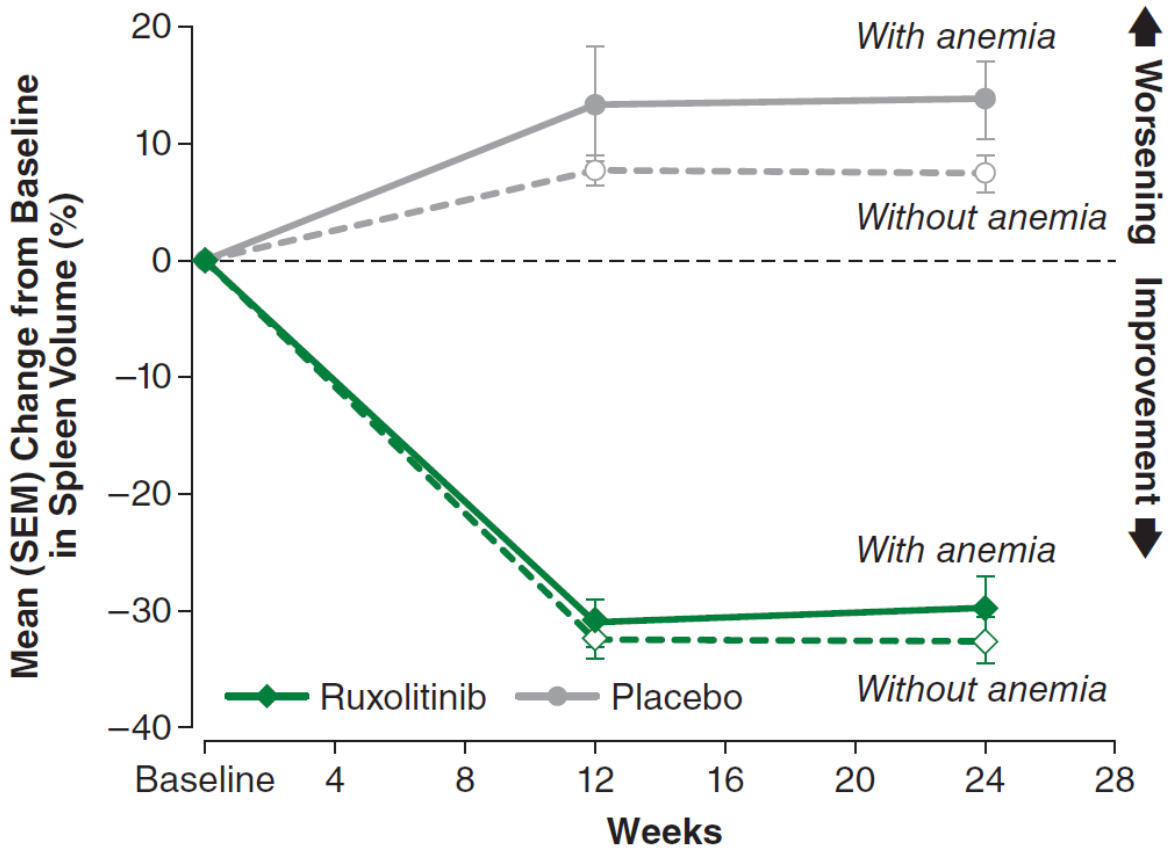
Ruxolitinib-treated patients with new-onset grade 3 or 4 anemia experienced improvements in symptoms that were similar to ruxolitinib-treated patients without anemia.



SEM denotes standard error of the mean.

Supplementary Figure S9D. Effect of Grade 3 or 4 Anemia on Spleen Volume Changes Over Time.

Spleen volume decreased in ruxolitinib-treated patients compared with placebo-treated patients, regardless of the presence or absence of grade 3 or 4 anemia.



SEM denotes standard error of the mean.

Supplementary Table S4. Red Blood Cell Transfusions.

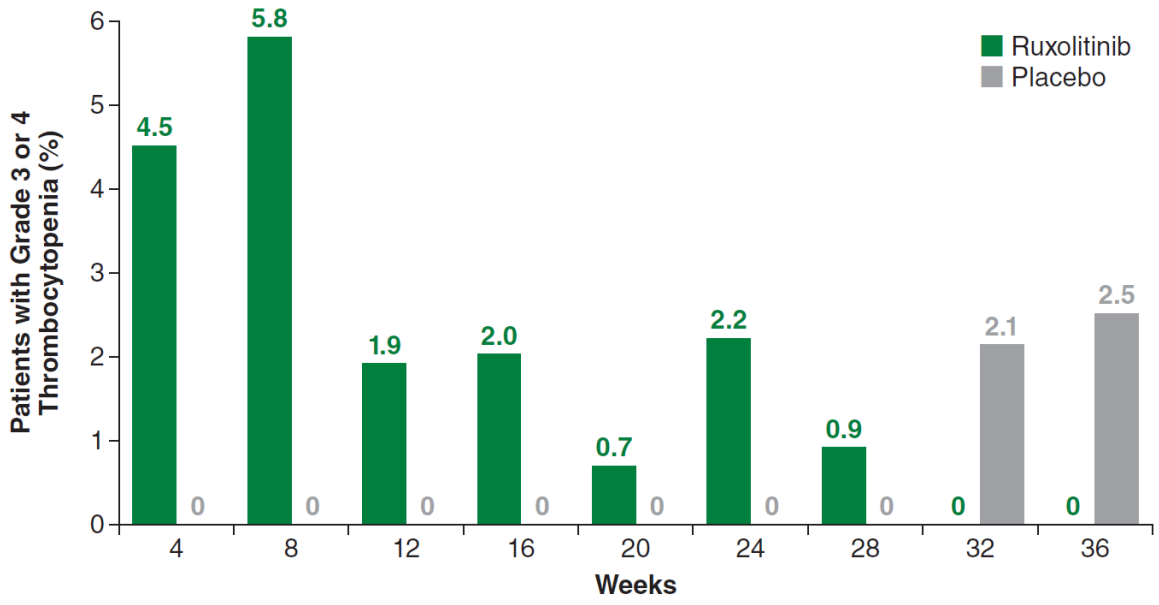
Using International Working Group for Myelofibrosis Research and Treatment response criteria, 41.2% of ruxolitinib-treated and 46.9% of placebo-treated patients who were transfusion-dependent at baseline changed their classification to transfusion-independent on study.

	Ruxolitinib	Placebo
Newly transfusion-independent by IWG criteria, N (%)*	14 (41)	15 (47)
Mean number of transfusions per month in patients transfused during randomized treatment	1.7	2.2

*Patients receiving at least 2 units packed red blood cells during the 4 weeks prior to randomization and no transfusions for at least 8 weeks while receiving randomized treatment. IWG denotes International Working Group.

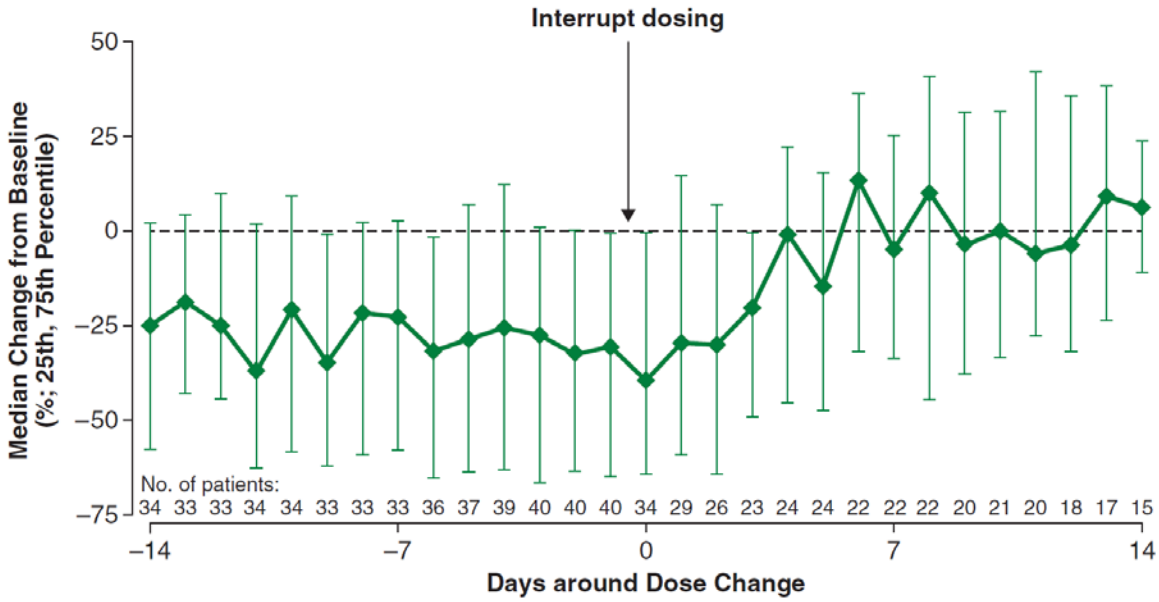
Supplementary Figure S10. Monthly Prevalence of New-Onset or Continuing Grade 3 or 4 Thrombocytopenia.

The monthly prevalence of grade 3 and 4 thrombocytopenia increased in the first 4 to 8 weeks of ruxolitinib treatment and then decreased to placebo levels.



Supplementary Figure S11. Total Symptom Score Before and During Dose Interruption.

For this analysis, the Total Symptom Score (TSS) around dose interruption represents the median for the 14 days before and after the first dose interruption. The baseline score is the 7-day moving average prior to study day 1. The graph represents the median percent change from baseline plus and minus 14 days around the first dose interruption in ruxolitinib-treated patients. Patients were counted only for those days around dose interruption for which data were available. Therefore, the number of patients for each time point differs in this analysis. Patients who had an interruption in ruxolitinib dosing experienced a gradual increase in TSS, which returned to baseline levels. Upon reinitiation of therapy, TSS improved again, consistent with levels prior to dose interruption (data not shown).



Supplementary Table S5. Adverse Events (Grade ≥3) During Dose Interruption*

	Ruxolitinib	Placebo
Patients with dose interruption, n / N	**49 / 155**	**54 / 151**
Total AEs with dose interruption, n (%)	**18 (36.7)**	**11 (20.4)**
Total grade ≥3 AEs with dose interruption, n (%)	**8 (16.3)**	**7 (13.0)**
Anemia	4 (8.2)	1 (1.9)
Thrombocytopenia	1 (2.0)	1 (1.9)
Acute renal failure	1 (2.0)	1 (1.9)
Disseminated intravascular coagulation	1 (2.0)	0
Abdominal pain	1 (2.0)	0
Gastrointestinal hemorrhage	1 (2.0)	0
Fatigue	1 (2.0)	0
Delirium	1 (2.0)	0
Splenic infarction	0	1 (1.9)
Atrial fibrillation	0	1 (1.9)
Ventricular dysfunction	0	1 (1.9)
Ascites	0	1 (1.9)
Gastric varices	0	1 (1.9)
Asthenia	0	1 (1.9)
Hyperbilirubinemia	0	1 (1.9)
Gout	0	1 (1.9)
Hepatic encephalopathy	0	1 (1.9)
Hydronephrosis	0	1 (1.9)

* Numbers reported are percentages of those who had a treatment interruption (not the total study population).

Supplementary Table S6. Adverse Events (Grade ≥3) After Discontinuation*

	Ruxolitinib	Placebo
Patients who discontinued, n / N	**21 / 155**	**37 / 151**
Total AEs with dose discontinuation, n (%)	**16 (76.1)**	**24 (64.8)**
Total grade ≥3 AEs with study discontinuation, n (%)	**12 (57.1)**	**17 (45.9)**
Thrombocytopenia	3 (14.3)	1 (2.7)
Pneumonia	2 (9.5)	4 (10.8)
AML	2 (9.5)	0
Fatigue	1 (4.8)	3 (8.1)
Renal failure	1 (4.8)	2 (5.4)
Subdural hematoma	1 (4.8)	1 (2.7)
Malnutrition	1 (4.8)	1 (2.7)
Sepsis	1 (4.8)	1 (2.7)
Septic shock	1 (4.8)	0
Pyrexia	1 (4.8)	0
Clostridial infection	1 (4.8)	0
Splenic hemorrhage	1 (4.8)	0
Splenic infarction	1 (4.8)	0
Platelet count increased	1 (4.8)	0
Hypokalemia	1 (4.8)	0
Muscular weakness	1 (4.8)	0
Respiratory failure	1 (4.8)	0
Abdominal pain	0	4 (10.8)
Disease progression	0	2 (5.4)
Dehydration	0	2 (5.4)
Hyponatremia	0	2 (5.4)
Hypoxia	0	2 (5.4)
Pulmonary embolism	0	2 (5.4)
Dyspnea	0	2 (5.4)
Febrile neutropenia	0	1 (2.7)
Leukocytosis	0	1 (2.7)
Splenomegaly	0	1 (2.7)
Cardiac failure	0	1 (2.7)
Tachycardia	0	1 (2.7)
Colitis	0	1 (2.7)
GI hemorrhage	0	1 (2.7)
Intestinal ischemia	0	1 (2.7)
Multiorgan failure	0	1 (2.7)
Staphylococcal infection	0	1 (2.7)
UTI	0	1 (2.7)

Fall	0	1 (2.7)
Splenic hematoma	0	1 (2.7)
Blood amylase increased	0	1 (2.7)
Blood magnesium increased	0	1 (2.7)
Lipase increased	0	1 (2.7)
Weight increased	0	1 (2.7)
Arthralgia	0	1 (2.7)
Musculoskeletal pain	0	1 (2.7)
Myelofibrosis	0	1 (2.7)
Loss of consciousness	0	1 (2.7)
Agitation	0	1 (2.7)
COPD	0	1 (2.7)
Pulmonary edema	0	1 (2.7)

* Numbers reported are percentages of those who discontinued (not the total study population).