CLINICAL SCIENCE

Filgotinib versus placebo or adalimumab in patients
with rheumatoid arthritis and inadequate response to
methotrexate: a phase III randomised clinical trial
Bernard Combe ‍ ‍,1 Alan Kivitz,2 Yoshiya Tanaka ‍ ‍,3 Désirée van der Heijde ‍ ‍,4
J Abraham Simon,5 Herbert S B Baraf,6 Uma Kumar ‍ ‍,7 Franziska Matzkies,8
Beatrix Bartok,8 Lei Ye,8 Ying Guo,8 Chantal Tasset,9 John S Sundy,8,10
Angelika Jahreis,8 Mark C Genovese,8 Neelufar Mozaffarian,11
Robert B M Landewé ‍ ‍,12 Sang-­Cheol Bae ‍ ‍,13 Edward C Keystone,14
Peter Nash ‍ ‍15
ABSTRACT
Objective To evaluate the efficacy and safety of the
Janus kinase-1-­preferential inhibitor filgotinib versus
►► Additional material is
placebo or tumour necrosis factor-α inhibitor therapy
published online only. To view,
please visit the journal online
in patients with active rheumatoid arthritis (RA) despite
(http://d​ x.​doi.o​ rg/​10.​1136/​
ongoing treatment with methotrexate (MTX).
annrheumdis-​2020-​219214).
Methods This 52-­week, multicentre, double-­blind,
placebo-­controlled and active-­controlled phase III
For numbered affiliations see
trial evaluated once-­daily oral filgotinib in patients
end of article.
with RA randomised 3:3:2:3 to filgotinib 200 mg
Correspondence to
(FIL200) or filgotinib 100 mg (FIL100), subcutaneous
Prof Bernard Combe,
adalimumab 40 mg biweekly, or placebo (through
Rheumatology, CHU Montpellier,
week 24), all with stable weekly background MTX.
34295 Montpellier, France;
​bernard.​combe@u​ montpellier.​fr The primary endpoint was the proportion of patients
achieving 20% improvement in American College of
Results from interim data cuts
Rheumatology criteria (ACR20) at week 12. Additional
were presented at the 2019
efficacy outcomes were assessed sequentially. Safety
Annual European Congress of
was assessed from adverse events and laboratory
Rheumatology, Madrid, Spain
abnormalities.
(Combe et al, Ann Rheum Dis.
2019; 78(Suppl 2):77–8) and at Results The proportion of patients (n=1755
the 2019 American College of
randomised and treated) achieving ACR20 at week 12
Rheumatology Annual Meeting,
was significantly higher for FIL200 (76.6%) and FIL100
Atlanta, Georgia (Combe et
(69.8%) versus placebo (49.9%; treatment difference
al, Arthritis Rheumatol. 2019;
(95% CI), 26.7% (20.6% to 32.8%) and 19.9% (13.6%
71(Suppl 10):A506). The FINCH
1 data were presented virtually
to 26.2%), respectively; both p<0.001). Filgotinib was
at the 2020 Annual European
superior to placebo in key secondary endpoints assessing
Congress of Rheumatology
RA signs and symptoms, physical function and structural
(Combe et al, Ann Rheum Dis.
damage. FIL200 was non-­inferior to adalimumab in
terms of Disease Activity Score in 28 joints with C
Received 30 September 2020
reactive protein ≤3.2 at week 12 (p<0.001); FIL100
Revised 5 January 2021
did not achieve non-­inferiority. Adverse events and
Accepted 6 January 2021
laboratory abnormalities were comparable among active
Published Online First
treatment arms.
27 January 2021
Conclusions Filgotinib improved RA signs and
symptoms, improved physical function, inhibited
radiographic progression and was well tolerated in
patients with RA with inadequate response to MTX.
© Author(s) (or their
FIL200 was non-­inferior to adalimumab.
employer(s)) 2021. Re-­use
Trial registration number NCT02889796.
permitted under CC BY-­NC. No
Handling editor Josef S
Smolen

commercial re-­use. See rights
and permissions. Published
by BMJ.
To cite: Combe B,
Kivitz A, Tanaka Y,
et al. Ann Rheum Dis
2021;80:848–858.

848  

INTRODUCTION

Scientific innovations have changed the landscape
of rheumatoid arthritis (RA) treatment. The cornerstone of RA treatment remains disease-­modifying

Key messages
What is already known about this subject?

►► Methotrexate (MTX) is the recommended initial

treatment for rheumatoid arthritis, with tumour
necrosis factor α inhibitors (TNFαi) as common
second-­line therapy in patients with inadequate
response.
►► Oral therapies that match or exceed TNFαi
efficacy in this population are still needed.
►► Filgotinib—a once-­daily, oral, Janus kinase1-­preferential inhibitor—with or without
MTX is superior relative to placebo treatment
in patients with rheumatoid arthritis with
inadequate response to MTX or prior biologic
failure.
What does this study add?

►► This is the first study to evaluate filgotinib

compared with TNFαi standard therapy
or placebo with stable background MTX
in patients with rheumatoid arthritis with
inadequate response to MTX but without prior
biologic failure, and to include a radiographic
endpoint.
►► Filgotinib treatment reduced rheumatoid
arthritis signs and symptoms, improved physical
function, inhibited radiographic progression and
appeared well tolerated for up to 52 weeks in
this population.
How might this impact on clinical practice or
future developments?
►► Filgotinib with background MTX could be
considered a treatment option in patients with
rheumatoid arthritis with inadequate response
to MTX.

antirheumatic drugs (DMARDs), including conventional synthetic DMARDs (csDMARDs), of which
methotrexate (MTX) is the gold standard, and
biologic DMARDs (bDMARDs) such as those
targeting cytokines (eg, tumour necrosis factor α
(TNFα), interleukin 6 or interleukin 1) and B or T


Rheumatoid arthritis


cells. Availability of TNFα inhibitors (TNFαi) in the late 1990s, non-TNFαi biologics in the 2000s and recently the targeted synthetic DMARDs has helped to reduce disease severity in patients with RA. Advances in RA management have further improved patient outcomes by focusing on treat-to-target strategies, pain and inflammation reduction, and administration convenience, in addition to efficacy and safety.3 Despite this focus, many patients do not achieve long-term responses with currently available therapies; in one study, only 10%–21% of patients initiating csDMARDs and 12%–24% initiating TNFαi therapy achieved remission within 12 months.4 Potential innovations that may further improve patient outcomes in RA include new oral therapies that perform as well as, or better than, existing standard of care (SOC), particularly in patients with intolerance or inadequate response to bDMARDs (bDMARD-IR).

The FINCH phase 3 programme was developed to study filgotinib, a Janus-associated kinase (JAK)-1-preferential inhibitor, for RA treatment. In FINCH 2, filgotinib significantly improved efficacy versus placebo in bDMARD-IR patients with active RA.5 FINCH 3 examined filgotinib use in patients with MTX-naïve RA. To address the MTX-IR population, the FINCH 1 study examined filgotinib versus placebo or adalimumab, all with background MTX, in MTX-IR patients with active RA.

METHODS

Study design and conduct

This randomised, double-blind, 52-week, placebo-controlled and active-controlled phase III trial was conducted at 303 sites in 30 countries from 30 August 2016 to 20 June 2019. The protocol and statistical analysis plan are provided in online supplemental files 1–3. All patients provided written informed consent. An independent data monitoring committee reviewed safety data periodically. An independent adjudication committee periodically reviewed all potential major cardiovascular adverse events (MACE) and thromboembolic events.

Study participants

Eligible patients were ≥18 years old at the time of consent and met the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism criteria for RA diagnosis.6 Patients had active moderate-to-severe RA, defined as ≥6 swollen joints and ≥6 tender joints (both at screening and on day 1 despite ongoing MTX treatment for ≥12 weeks and stable at 7.5–25 mg/week for ≥4 weeks). Additional inclusion criteria were seropositivity for anticyclic citrullinated peptide (anti-CCP) antibodies or rheumatoid factor (RF); ≥1 joint erosion on hand/wrist and foot radiographs, or ≥3 erosions if negative for RF and anti-CCP; or serum C reactive protein (CRP) ≥6 mg/L. Key exclusion criteria included previous use of JAK inhibitors (JAKi) or adalimumab, prior non-response or intolerance to any bDMARD, and recent use of csDMARDs other than MTX or stably dosed hydroxychloroquine or chloroquine; concomitant, stably dosed non-steroidal anti-inflammatory drugs or glucocorticoids (≤10 mg/day prednisone/equivalent) were permitted.

Interventions

Eligible patients were randomly assigned (3:3:2:3) to oral filgotinib 200 mg (FIL200) or filgotinib 100 mg (FIL100) once daily, subcutaneous adalimumab 40 mg every 2 weeks, or placebo, all with stable background MTX; other concomitant medications were to be kept stable as much as possible. Study participants were blinded to treatment and received placebo tablets matching FIL200 and/or FIL100; patients not assigned to active adalimumab received matching placebo injections. At week 24, placebo-treated patients were rerandomised (1:1) to FIL200 or FIL100 and continued background MTX. Per protocol, patients without adequate treatment response (<20% improvement from baseline in either swollen joint count 66 or tender joint count 68) at week 14 or two consecutive visits after week 30 discontinued study treatment but continued study visits, using investigator-specified SOC RA therapy.

Endpoints and assessments

The primary efficacy endpoint was ACR20 response (20% improvement in ACR criteria)7 at week 12. Key secondary efficacy endpoints tested hierarchically at week 12 (unless otherwise specified) were change from baseline score on the Health Assessment Questionnaire-Disability Index (HAQ-DI),8 9 proportion of patients with Disease Activity Score in 28 joints with CRP (DAS28(CRP)) <2.6,10 change from baseline van der Heijde modified total Sharp score (mTSS)11 at week 24 (radiographic assessment details in online supplemental methods), non-inferiority of filgotinib versus adalimumab for a proportion of patients with DAS28(CRP) ≤3.2, change from baseline Short Form-36 Physical Component Summary12 and Functional Assessment of Chronic Illness Therapy-Fatigue score,13 superiority of filgotinib versus adalimumab for a proportion of patients with DAS28(CRP) ≤3.2, non-inferiority of filgotinib versus adalimumab for a proportion of patients with DAS28(CRP) <2.6, and superiority of filgotinib versus adalimumab for a proportion of patients with DAS28(CRP) ≥2.6. Other secondary endpoints included ACR50/70; low disease activity defined as Clinical Disease Activity Index (CDAI) ≤10 or Simplified Disease Activity Index (SDAI) ≤11;14 and remission defined as CDAI ≤2.8, SDAI ≤3.3 or Boolean remission.15 Safety was assessed from laboratory tests and adverse events (AEs). Positively adjudicated MACE and thromboembolic events were reported.

Statistical analysis

A sample size of 450 patients per filgotinib and placebo group was estimated to provide >90% power at a two-sided α of 0.05 to test the superiority of FIL200 versus placebo for change from baseline mTSS at week 24, based on other RA studies with radiography.16–18 This sample size also provided >95% power to detect a 20% difference in ACR20 for filgotinib versus placebo. Assuming similar DAS28(CRP) ≤3.2 response rates for filgotinib and adalimumab, approximately 300 adalimumab-treated patients were required to ensure >90% power at a two-sided α of 0.05 to demonstrate non-inferiority of FIL200 versus adalimumab. Consistent with regulatory guidance, non-inferiority assessments were based on the method of Liu et al.,19 which does not require a prespecified fixed non-inferiority margin or constancy and assay sensitivity assumptions.20 Non-inferiority testing assessed whether the effect of each filgotinib dose (response rate difference between filgotinib and placebo) preserves >50% of the effect of adalimumab (difference in response rate between adalimumab and placebo). The 50% non-inferiority margin of DAS28(CRP) ≤3.2 and ≤2.6 at weeks 12 and 24 based on FINCH 1 data are presented in online supplemental table S1.

Type I error rate was controlled by hierarchical testing of primary and key secondary endpoints at a two-sided α of 0.05 (online supplemental figure S1). The primary analysis tested the superiority of FIL200 versus placebo for ACR20 at week 12 using a logistic regression model, with treatment and stratification factors included as covariates. Hypothesis testing for
Rheumatoid arthritis

key secondary endpoints commenced only after the primary endpoint reached statistical significance and proceeded sequentially until a null hypothesis was not rejected, after which exploratory p values are reported for the remaining hypotheses.

All analyses were based on data from patients who received ≥1 dose of study drug. For binary endpoints, a logistic regression model including treatment and stratification factors (geographical region, prior exposure to bDMARDs, and RF or CCP antibody positivity at screening) was used. Treatment effect on continuous endpoint change from baseline was evaluated using a mixed-effects model for repeated measures, with treatment, visit, treatment by visit interaction, stratification factors and baseline value included as fixed effects and subject as a random effect. Patients who required rescue therapy or had missing values were defined as non-responders, and non-responder imputation (NRI) was employed for primary and key secondary binary endpoint analyses. Multiple imputation was conducted to determine the impact of NRI on the robustness of results (online supplemental methods and table S2). Safety analyses of AEs and laboratory data were summarised by treatment group using descriptive statistics.

Patient and public involvement
Patients and the public were not involved in the design, conduct, reporting or dissemination of this research.

RESULTS

Study participants
A total of 1755 patients received study treatment (enrolment by country; online supplemental figure S2), and 87.4% completed the study visits through the 24-week placebo-controlled period. The reasons for discontinuation are summarised in figure 1. At week 14, 4.8% of FIL200-treated, 6.0% of FIL100-treated, 4.0% of adalimumab-treated and 8.6% of placebo-treated patients had inadequate response to treatment and were mandated to SOC. After week 24, four patients receiving FIL200, three receiving FIL100, three receiving adalimumab and two in each placebo-to-filotitinib arm discontinued study drug due to lack of efficacy. Baseline demographics, concomitant medications and disease characteristics were similar among the treatment arms (table 1).

Efficacy
ACR20 responses at week 12 were significantly greater in patients receiving filgotinib versus placebo: 76.6% for FIL200 and 69.8% for FIL100 vs 49.9% for placebo (all p<0.001) (table 2, figure 2A). Significant improvements at week 12 with filgotinib versus placebo treatment were also observed in key secondary endpoints, including HAQ-DI and DAS28(CRP) <2.6 (all p<0.001) (table 2). Radiographic progression of structural joint damage was significantly reduced in both filgotinib dose arms versus placebo at week 24 (p<0.001 for FIL200; p<0.001 for FIL100) (figure 3). FIL200 was non-inferior to adalimumab at week 12 for DAS28(CRP) ≤3.2 (p<0.001); FIL100 did not

Figure 1  Patient disposition. *23 (4.8%) patients treated with filgotinib 200mg, 29 (6.0%) patients treated with filgotinib 100mg, 13 (4.0%) patients treated with adalimumab, and 41 (8.6%) patients treated with placebo did not have adequate response to treatment per protocol at week 14. †3 (0.7%) patients treated with filgotinib 200mg, 2 (0.5%) patients treated with filgotinib 100mg, 3 (1.0%) patients treated with adalimumab, 0 patient treated with placebo and rerandomised to filgotinib 200mg at week 24, and 4 (2.2%) patients treated with placebo and rerandomised to filgotinib 100mg at week 24 failed to maintain response to treatment per protocol after week 30. ADA, adalimumab; FIL, filgotinib; PBO, placebo; W, week.
**Table 1** Baseline demographics and disease characteristics

<table>
<thead>
<tr>
<th></th>
<th>FIL200 (n=475)</th>
<th>FIL100 (n=480)</th>
<th>ADA (n=325)</th>
<th>PBO (n=475)</th>
<th>Total (N=1755)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex at birth</strong>, n (%), female</td>
<td>379 (79.8)</td>
<td>399 (83.1)</td>
<td>266 (81.8)</td>
<td>391 (82.3)</td>
<td>1435 (81.8)</td>
</tr>
<tr>
<td><strong>Age, years</strong></td>
<td>52±12.8</td>
<td>53±12.6</td>
<td>53±12.9</td>
<td>53±12.8</td>
<td>53±12.7</td>
</tr>
<tr>
<td><strong>Weight, kg</strong></td>
<td>70.6±17.5</td>
<td>69.9±16.9</td>
<td>71.5±17.4</td>
<td>70.6±16.8</td>
<td>70.6±17.1</td>
</tr>
<tr>
<td><strong>Body mass index, kg/m²</strong></td>
<td>26.7±5.7</td>
<td>26.4±5.8</td>
<td>26.9±6.0</td>
<td>27.0±5.9</td>
<td>26.7±5.8</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td>312 (65.7)</td>
<td>324 (67.5)</td>
<td>229 (70.5)</td>
<td>319 (67.2)</td>
<td>1184 (67.5)</td>
</tr>
<tr>
<td>Asian</td>
<td>122 (25.7)</td>
<td>115 (24.0)</td>
<td>70 (21.6)</td>
<td>70 (14.7)</td>
<td>411 (23.4)</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>27 (5.7)</td>
<td>27 (5.6)</td>
<td>10 (3.1)</td>
<td>12 (2.5)</td>
<td>50 (2.9)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>6 (1.3)</td>
<td>7 (1.5)</td>
<td>10 (3.1)</td>
<td>12 (2.5)</td>
<td>35 (2.0)</td>
</tr>
<tr>
<td>Other*</td>
<td>8 (1.7)</td>
<td>6 (1.3)</td>
<td>1 (0.3)</td>
<td>5 (1.1)</td>
<td>20 (1.1)</td>
</tr>
<tr>
<td>Not permitted</td>
<td>0</td>
<td>1 (0.2)</td>
<td>0</td>
<td>1 (0.2)</td>
<td>2 (0.1)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td>404 (85.1)</td>
<td>399 (83.1)</td>
<td>268 (82.5)</td>
<td>400 (84.2)</td>
<td>1471 (83.8)</td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>8.8±21.0</td>
<td>9.0±23.0</td>
<td>196 (70.5)</td>
<td>274 (57.7)</td>
<td>1064 (60.6)</td>
</tr>
<tr>
<td>Median, years</td>
<td>7.3±7.2</td>
<td>8.3±8.2</td>
<td>8.0±7.4</td>
<td>7.8±7.6</td>
<td>7.8±7.6</td>
</tr>
<tr>
<td><strong>hsCRP, mg/L</strong></td>
<td>16±1.0</td>
<td>16.7±2.0</td>
<td>19.6±2.8</td>
<td>17.6±2.6</td>
<td>18.8±2.6</td>
</tr>
<tr>
<td>Median</td>
<td>12.0±2.0</td>
<td>12.5±2.5</td>
<td>12.5±2.5</td>
<td>11.5±2.5</td>
<td>12.0±2.5</td>
</tr>
<tr>
<td><strong>Erosion score &gt;0, n (%)</strong></td>
<td>399 (84.0)</td>
<td>411 (85.1)</td>
<td>277 (82.5)</td>
<td>400 (84.2)</td>
<td>1471 (83.8)</td>
</tr>
<tr>
<td><strong>JSN score</strong></td>
<td>15±8.5</td>
<td>15±8.5</td>
<td>16±8.4</td>
<td>16±8.5</td>
<td>16±8.5</td>
</tr>
<tr>
<td><strong>bDMARD-naive, n (%)</strong></td>
<td>458 (96.4)</td>
<td>464 (96.7)</td>
<td>317 (97.5)</td>
<td>466 (98.7)</td>
<td>1708 (97.3)</td>
</tr>
<tr>
<td><strong>MTX dose, mg/week</strong></td>
<td>15.3±4.9</td>
<td>15.5±4.8</td>
<td>15.4±4.8</td>
<td>14.9±4.5</td>
<td>15.3±4.8</td>
</tr>
<tr>
<td><strong>Concurrent oral steroids, n (%)</strong></td>
<td>229 (48.2)</td>
<td>229 (47.7)</td>
<td>140 (43.1)</td>
<td>217 (45.7)</td>
<td>815 (46.4)</td>
</tr>
<tr>
<td><strong>≤5 mg/day, n (%)</strong></td>
<td>152 (10.6)</td>
<td>160 (10.5)</td>
<td>96 (6.8)</td>
<td>152 (70.0)</td>
<td>560 (60.7)</td>
</tr>
<tr>
<td><strong>Steroid dose, mg/day</strong></td>
<td>6.2±3.4</td>
<td>6.1±2.5</td>
<td>5.9±2.2</td>
<td>6.0±2.8</td>
<td>6.0±2.8</td>
</tr>
<tr>
<td><strong>Concurrent antimalarials, n (%)</strong></td>
<td>64 (13.5)</td>
<td>59 (12.3)</td>
<td>39 (12.0)</td>
<td>63 (13.3)</td>
<td>225 (12.8)</td>
</tr>
<tr>
<td><strong>VAS, mm</strong></td>
<td>67±19.2</td>
<td>65±19.7</td>
<td>67±19.1</td>
<td>68±19.3</td>
<td>67±19.2</td>
</tr>
<tr>
<td><strong>Pain, VAS, mm</strong></td>
<td>66±16.0</td>
<td>65±16.2</td>
<td>67±15.5</td>
<td>66±16.2</td>
<td>65±16.2</td>
</tr>
<tr>
<td><strong>HAQ-DI</strong></td>
<td>1.6±0.6</td>
<td>1.6±0.6</td>
<td>1.6±0.6</td>
<td>1.6±0.6</td>
<td>1.6±0.6</td>
</tr>
<tr>
<td><strong>SF-36 PCS</strong></td>
<td>33.4±7.2</td>
<td>33.6±7.8</td>
<td>32.8±7.7</td>
<td>32.9±7.1</td>
<td>33.2±7.4</td>
</tr>
<tr>
<td><strong>SF-36 MCS</strong></td>
<td>43.9±10.4</td>
<td>44.6±10.4</td>
<td>44.1±10.4</td>
<td>43.4±10.0</td>
<td>43.6±10.6</td>
</tr>
<tr>
<td><strong>Facit-F</strong></td>
<td>27.6±10.7</td>
<td>27.8±10.6</td>
<td>27.2±10.2</td>
<td>26.9±10.3</td>
<td>27.4±10.5</td>
</tr>
</tbody>
</table>

Values are mean±SD.

*Includes patients recorded as Native Hawaiian/Pacific Islander and ‘Other’. Race was not recorded for one patient receiving FIL100 and one patient receiving PBO due to local regulations.

†n=1 missing.

‡n=2 missing.

§Campaign A: FIL200, n=467; FIL100, n=471; ADA, n=319; PBO, n=466.

¶Campaign A: FIL200, n=8 missing; FIL100, n=9 missing; ADA, n=6 missing; PBO, n=9 missing.

**FIL100, n=479; ADA, n=324.

††Percent of patients with concurrent oral corticosteroid use on first dosing date.

†‡FIL200, n=226; FIL100, n=229; ADA, n=140; PBO, n=217.

§§FIL200, n=473; FIL100, n=479; ADA, n=323; PBO, n=474.

¶¶FIL200, n=472; FIL100, n=477; ADA, n=319; PBO, n=469.

ADA, adalimumab; anti-CCP, anticyclic citrullinated protein antibody; bDMARD, biologic disease-modifying antirheumatic drug; CDAI, Clinical Disease Activity Index; DAS28(CRP), Disease Activity Score in 28 joints with C reactive protein; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; HAQ-DI, Health Assessment Questionnaire-Disability Index; hsCRP, high-sensitivity C reactive protein; JSN, joint space narrowing; MCS, Mental Component Summary; mTSS, van der Heijde modified total Sharp score; MTX, methotrexate; PBO, placebo; PCS, Physical Component Summary; PGA, Physician’s Global Assessment; Q1, first quartile; Q3, third quartile; RA, rheumatoid arthritis; RF, rheumatoid factor; SDAI, Simplified Disease Activity Index; SF-36, Short Form-36; SGA, Subject’s Global Assessment; SJC66, swollen joint count of 66 joints; TJC68, tender joint count of 68 joints; VAS, visual analogue scale.
Rheumatoid arthritis

Table 2  Primary and key secondary efficacy outcomes during the placebo-controlled period*

<table>
<thead>
<tr>
<th></th>
<th>FIL200 (n=475)</th>
<th>FIL100 (n=480)</th>
<th>ADA (n=325)</th>
<th>PBO (n=475)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20, week 12</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>364/475</td>
<td>325/480</td>
<td>229/325</td>
<td>237/475</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>26.7 (20.6 to 32.8)</td>
<td>19.9 (13.6 to 26.2)</td>
<td>20.6 (13.6 to 27.5)</td>
<td></td>
</tr>
<tr>
<td>P value vs placebo</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Key secondary outcomes with hierarchical testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>457</td>
<td>459</td>
<td>311</td>
<td>435</td>
</tr>
<tr>
<td>Means±SD</td>
<td>−0.69±0.61</td>
<td>−0.56±0.56</td>
<td>−0.61±0.56</td>
<td>−0.42±0.54</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>−0.29 (−0.36 to −0.22)</td>
<td>−0.17 (−0.24 to −0.10)</td>
<td>−0.20 (−0.28 to −0.13)</td>
<td></td>
</tr>
<tr>
<td>P value vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DAS28(CRP) &lt;2.6, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n/N</td>
<td>162/475</td>
<td>114/480</td>
<td>77/325</td>
<td>44/475</td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>34.1 (29.7 to 38.5)</td>
<td>23.8 (19.8 to 27.7)</td>
<td>23.7 (18.9 to 28.5)</td>
<td>9.3 (6.6 to 12.0)</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>24.8 (19.6 to 30.0)</td>
<td>14.5 (9.7 to 19.3)</td>
<td>14.4 (8.9 to 20.0)</td>
<td></td>
</tr>
<tr>
<td>P value vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>mTSS change from baseline to week 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>405</td>
<td>404</td>
<td>271</td>
<td>351</td>
</tr>
<tr>
<td>Means±SD</td>
<td>0.13±0.9</td>
<td>0.17±0.91</td>
<td>0.16±0.95</td>
<td>0.37±1.42</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>−0.27 (−0.43 to −0.12)</td>
<td>−0.25 (−0.40 to −0.10)</td>
<td>−0.22 (−0.39 to −0.05)</td>
<td></td>
</tr>
<tr>
<td>P value vs PBO</td>
<td>&lt;0.001</td>
<td>0.012†</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-inferiority DAS28(CRP) ≤3.2, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n/N</td>
<td>236/475</td>
<td>186/480</td>
<td>141/325</td>
<td>111/475</td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>49.7 (45.1 to 54.3)</td>
<td>38.8 (34.3 to 43.2)</td>
<td>43.4 (37.8 to 48.9)</td>
<td>23.4 (19.5 to 27.3)</td>
</tr>
<tr>
<td>P value vs ADA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Key secondary outcomes without multiplicity adjustment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 PCS change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>459</td>
<td>463</td>
<td>310</td>
<td>440</td>
</tr>
<tr>
<td>Means±SD</td>
<td>9.2±8.1</td>
<td>8.5±7.7</td>
<td>8.4±7.9</td>
<td>5.8±7.1</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>3.7 (2.8 to 4.6)</td>
<td>3.1 (2.2 to 4.0)</td>
<td>2.6 (1.6 to 3.6)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p value vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FACIT-F change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>452</td>
<td>455</td>
<td>304</td>
<td>432</td>
</tr>
<tr>
<td>Means±SD</td>
<td>9.2±9.8</td>
<td>9.1±10.2</td>
<td>8.8±9.2</td>
<td>6.8±9.9</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)†</td>
<td>2.8 (1.7 to 3.9)</td>
<td>2.6 (1.5 to 3.7)</td>
<td>2.1 (0.9 to 3.3)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p value vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Superiority DAS28(CRP) ≤3.2, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference vs ADA (95% CI)†</td>
<td>6.3 (−1.0 to 13.6)</td>
<td>−4.6 (−11.8 to 2.6)</td>
<td>0.069</td>
<td>0.18</td>
</tr>
<tr>
<td>Exploratory p value vs ADA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-inferiority DAS28(CRP) ≤2.6, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p value vs ADA</td>
<td>&lt;0.001</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superiority DAS28(CRP) ≤2.6, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference vs ADA (95% CI)†</td>
<td>10.4 (3.9 to 17.0)</td>
<td>0.1 (−6.2 to 6.3)</td>
<td>0.001</td>
<td>0.99</td>
</tr>
</tbody>
</table>

*Hierarchical testing according to prespecified, US Food and Drug Administration-reviewed, statistical analysis plan. Patients who had missing values were defined as non-responders, and NRI was employed for both primary and key secondary analyses.
†Difference in response rates vs placebo or ADA for categorical outcomes; least-squares mean difference vs placebo or ADA for continuous outcomes.
‡Exploratory p value without multiplicity adjustment.
ACR20, American College of Rheumatology criteria 20% decrease from baseline; ADA, adalimumab; DAS28(CRP), Disease Activity Score in 28 joints with C reactive protein; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; HAQ-DI, Health Assessment Questionnaire-Disability Index; mTSS, van der Heijde modified total Sharp score; NRI, non-responder imputation; PBO, placebo; SF-36 PCS, Short Form 36 Physical Component Summary.

achieve non-inferiority versus adalimumab for this measure (p=0.054) (table 2).

The remaining key secondary endpoints were not adjusted for multiplicity and are presented as exploratory analyses (table 2). ACR50/70 responses at week 12 were higher following FIL200 (47.2%/26.1%), FIL100 (36.5%/18.5%) or adalimumab (35.1%/14.2%) compared with placebo (19.8%/6.7%) (figure 2B,C). Response rates for DAS28(CRP) ≤3.2 at week 12 were higher in both filgotinib dose arms and placebo (table 2). Patients receiving filgotinib achieved higher rates of

852
remission and low disease activity across several composite disease measures (DAS28(CRP), CDAI, SDAI, Boolean remission) versus placebo at weeks 12 and 24 (figure 4A,B). Filgotinib efficacy was sustained through week 52 (figures 2A–C and 4A,B, online supplemental tables S3 and S4, figure S3).

Changes from baseline in ACR and DAS28(CRP) components at week 12 were generally consistent with the primary and key secondary efficacy outcomes, although the effect of FIL versus adalimumab or placebo treatment was more pronounced for high-sensitivity CRP compared with other measures (online supplemental table S3). However, in post-hoc exploratory analyses, FIL200 was non-inferior to adalimumab for CDAI low disease activity and remission at weeks 12 and 24 (online supplemental table S3). In a subanalysis of proportion of patients achieving ACR20 at week 12 across countries, the placebo response rate ranged from 36.8% to 59.2% and was highest in group B (predominantly Eastern Europe) and group C (Mexico and Argentina) (online supplemental table S6).

Figure 2 Proportions of patients achieving (A) ACR20, (B) ACR50 and (C) ACR70 through week 52. Error bars show 95% CI. Additional statistical details are available in online supplemental table S3 and all response rates in online supplemental table S7. **p<0.01, ***p<0.001 versus PBO, not adjusted for multiplicity and should be considered exploratory except for ACR20 for FIL200 and FIL100 versus PBO at week 12. *p<0.05, **p<0.01, ***p<0.001 versus ADA, not adjusted for multiplicity and should be considered exploratory. ACR20/50/70, 20%/50%/70% improvement from baseline by the American College of Rheumatology core criteria; ADA, adalimumab; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; PBO, placebo.

Figure 3 Radiographic progression through week 24. (A) mTSS change from baseline, (B) erosion score change from baseline and (C) joint space narrowing change from baseline. Data from campaign A (through week 24) are shown. Supporting data are shown in online supplemental table S4. Patient numbers at each time point in (B) and (C) are the same as for (A). Error bars represent the SE of the LS mean. *p<0.05, **p<0.01, ***p<0.001 versus PBO, not adjusted for multiplicity and should be considered exploratory except for mTSS change from baseline following FIL200 and FIL100 versus PBO at week 24. Difference for mTSS change from baseline at week 24 following treatment with FIL200 or FIL100 versus ADA was explored and was not significant for either dose. ADA, adalimumab; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; LS, least-squares; mTSS, van der Heijde modified total Sharp score; PBO, placebo.
A

Figure 4 Proportions of patients achieving (A) low disease activity and (B) DAS28(CRP) <2.6 or remission at weeks 12, 24 and 52. Error bars show 95% CI. Additional statistical details are available in online supplemental table S3. *p<0.05, **p<0.01, ***p<0.001 versus placebo, not adjusted for multiplicity and should be considered exploratory except for FIL200 and FIL100 versus placebo for DAS28(CRP) <2.6 at week 12. Non-inferior versus adalimumab. *p<0.05, **p<0.01, ***p<0.001 versus ADA, not adjusted for multiplicity and should be considered exploratory. ADA, adalimumab; Boolean, Boolean remission; CDW, SDAI, Clinical Disease Activity Index; DAS28(CRP), Disease Activity Score in 28 joints with C reactive protein; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; PBO, placebo; SDAI, Simplified Disease Activity Index.

Safety

Treatment-emergent AEs (TEAEs) are presented in table 3. The incidence of serious TEAEs during the active-controlled period through week 52 was similar among all original active treatment arms and in patients randomised from placebo to filgotinib. During the placebo-controlled period, malignancy (excluding non-melanoma skin cancer) was reported in five patients: one (0.2%), one (0.3%) and three (0.6%) patients receiving FIL100, adalimumab and placebo, respectively. Venous thromboembolism (VTE) was reported in three patients: one (0.2%) receiving FIL200 and two (0.4%) receiving placebo. Adjudicated MACE occurred in four patients: one (0.2%) receiving FIL100, one (0.3%) receiving adalimumab and two (0.4%) receiving placebo. All patients with VTE and MACE had at least one risk factor, and no patient with deep vein thrombosis (DVT) or pulmonary embolism had a platelet count measurement above 600×10⁹/L.

Through week 24, death was reported in five patients: two (0.4%) receiving FIL200 (both attributed to septic shock), one (0.2%) receiving FIL100 (myocardial infarction) and two (0.4%) receiving placebo (one toxic reaction to amoxicillin/clavulanic acid and one non-TEAE septic shock). Four additional deaths occurred in the active-controlled period: one patient receiving FIL200 (alveolitis), one receiving adalimumab (sepsis), one placebo-treated patient randomised to FIL200 (acute DVT) and one placebo-treated patient randomised to FIL100 (primary varicella). Additional details of the DVT-associated and primary varicella-associated deaths are provided in online supplemental results.

Overall, infectious and serious infectious TEAEs occurred more frequently in patients receiving filgotinib or adalimumab versus placebo through week 24. Serious infections occurring in >2 patients were pneumonia (13 patients), cellulitis (3 patients) and bronchitis (3 patients). Through week 24, herpes zoster (excluding primary varicella) occurred in all treatment arms in 0.4% of patients receiving either filgotinib dose or placebo and in 0.6% of patients receiving adalimumab. Through week 52, serious infections occurred in 2.7%, 2.7% and 3.1% and herpes zoster occurred in 1.3%, 0.8% and 0.6% of patients receiving FIL200, FIL100 and adalimumab, respectively. In 14% of patients randomised in Asia (online supplemental figure S2), the frequency of herpes zoster was 1%, 3% and 0% for patients receiving FIL200, FIL100 and adalimumab, respectively. In week 52, and 2% in placebo-treated patients through week 24. Both reported opportunistic infections were in patients receiving adalimumab: one patient with Pneumocystis jirovecii pneumonia before week 24 and one patient with active Mycobacterium tuberculosis after week 24.

Grade 3/4 changes in laboratory values are shown in table 4. Mean haemoglobin levels were stable or increased across all treatment arms, with no imbalance in individual decreased haemoglobin events or grade 3 changes. Decreases in neutrophil and lymphocyte levels were seen in filgotinib-treated and adalimumab-treated patients. Grade ≥3 lymphopenia and neutropenia were more frequent in patients receiving filgotinib versus placebo. The majority of white cell count abnormalities were grade 1/2, not associated with infection, and resolved at follow-up testing. No grade ≥3 changes in platelet counts were observed. Higher mean creatinine levels were observed in patients receiving filgotinib versus adalimumab or placebo. Grade 3/4 serum creatinine elevations were reported in three patients: one receiving FIL100 and two receiving placebo, all before week 24. Mean creatine kinase and low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol levels were increased in patients treated with filgotinib versus placebo, without meaningful change in the LDL to HDL cholesterol ratio.

DISCUSSION

The FINCH 1 study assessed filgotinib, an oral JAK1-preferential inhibitor, to address the unmet needs for RA treatment in MTX-IR patients. Two doses of filgotinib were compared with adalimumab and placebo, all with background MTX. Both filgotinib doses were superior to placebo for ACR20 response and hierarchical key secondary endpoints evaluating signs and symptoms, physical function and structural damage. Although conclusions are limited for tests without multiplicity adjustment, proportions of patients achieving various measures of low disease activity and remission were generally consistent with DAS28(CRP) <2.6 and DAS28(CRP) ≤3.2 response results.

These phase III results confirm those of two phase II studies investigating filgotinib with or without MTX versus placebo in MTX-IR patients and a phase III study (FINCH 2) in bDMARD-refractory patients, and are consistent with the
Table 3  Treatment-emergent adverse events through week 24 and week 52

<table>
<thead>
<tr>
<th>Event</th>
<th>PBO-controlled period (weeks 0–24)</th>
<th>Weeks 0–52</th>
<th>PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FIL200 (n=475)</td>
<td>FIL100 (n=480)</td>
<td>ADA (n=325)</td>
</tr>
<tr>
<td>TEAEs, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any TEAE</td>
<td>287 (60.4)</td>
<td>287 (59.8)</td>
<td>186 (57.2)</td>
</tr>
<tr>
<td>TE SAE</td>
<td>21 (4.4)</td>
<td>24 (5.0)</td>
<td>14 (4.3)</td>
</tr>
<tr>
<td>TEAE leading to treatment discontinuation</td>
<td>15 (3.2)</td>
<td>9 (1.9)</td>
<td>13 (4.0)</td>
</tr>
<tr>
<td>Deaths</td>
<td>2 (0.4)</td>
<td>1 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Deaths in &gt;5% of patients*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Infectious AEs               | 133 (28.0)                         | 128 (26.7) | 88 (27.1)    | 105 (22.1)   | 206 (43.4)   | 194 (40.4)   | 129 (39.7)   | 45 (23.7)    | 39 (20.4)    |
| Serious infectious AEs       | 8 (1.7)                            | 8 (1.7)    | 8 (2.5)      | 4 (0.8)      | 13 (2.7)     | 13 (2.7)     | 10 (3.1)     | 1 (0.5)      | 2 (1.0)      |
| Malignancy                   |                                    |            |              |              |              |            |              |              |              |
| MACE†                        | 0                                  | 1 (0.2)    | 1 (0.3)      | 2 (0.4)      | 0            | 2 (0.4)     | 1 (0.3)     | 1 (0.5)      | 1 (0.5)      |
| MACE‡                        |                                    |            |              |              |              |            |              |              |              |

*TEAEs occurring in >5% of patients in a single treatment arm during either study period.
†Positively adjudicated.
‡Excluding MACE and VTE.

Results for other JAKis in MTX-IR patients with RA. FIL200 efficacy was statistically non-inferior to adalimumab for a proportion of patients with DAS28(CRP) ≤3.2 at week 12, a treat-to-target checkpoint, and remained non-inferior in exploratory analyses of CDAI low disease activity and remission, suggesting direct effects of JAK inhibition on high-sensitivity CRP. However, FIL200 treatment effect was inferior to adalimumab for patients with CDAI low disease activity and remission. The low frequency of herpes zoster does not appear attributable to geography; the proportion of FINCH 1 patients enrolled in Asian countries (14%) was comparable relative to similar JAKi studies (3%–18%). No cases of opportunistic infection or tuberculosis were observed in filgotinib-treated patients. Rates of AEs in filgotinib-treated patients were consistent with or below those from a recent meta-analysis on JAKi treatment in RA.

Filgotinib benefits must be considered in the context of risks. In this study, serious TEAEs and discontinuations due to TEAEs were similar among treatment arms through week 24. Safety data remained consistent over the entire 52-week study. Adjudicated MACE and VTE were observed in all treatment arms at frequencies similar to reported background rates in patients with RA.

Infections were increased in patients treated with FIL200 versus placebo, with similar rates of serious infections across active treatment groups. The frequency of herpes zoster was low and similar across all groups through week 24; the number of uncomplicated cases increased slightly after week 24 in the filgotinib versus adalimumab treatment arms. The low frequency of herpes zoster does not appear attributable to geography; the proportion of FINCH 1 patients enrolled in Asian countries (14%) was comparable relative to similar JAKi studies (3%–18%). No cases of opportunistic infection or tuberculosis were observed in filgotinib-treated patients. Rates of AEs in filgotinib-treated patients were consistent with or below those from a recent meta-analysis on JAKi treatment in RA.

Filgotinib was associated with decreases in neutrophil, lymphocyte and platelet counts and increases in lipid, creatine kinase and creatinine levels, as previously reported for filgotinib and other JAKis. There were small numerical differences in frequencies of grade 3/4 neutropaenia and lymphopaenia in patients treated with filgotinib versus placebo. Treatment with filgotinib was associated with small increases in fasting total, LDL and HDL cholesterol without affecting fasting LDL to HDL ratio, consistent with the hypothesis that JAKi treatment suppresses elevated cholesterol ester catabolism in patients with active RA and normalises their cholesterol levels towards the range in healthy volunteers.
Table 4  Laboratory values and grade ≥3 abnormalities through week 24 and week 52

<table>
<thead>
<tr>
<th></th>
<th>PBO-controlled period (weeks 0–24)</th>
<th>Weeks 0–52</th>
<th>PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FIL200 (n=475)</td>
<td>FIL100 (n=480)</td>
<td>ADA (n=325)</td>
</tr>
<tr>
<td>Haemoglobin, g/L</td>
<td>2.11 (1)</td>
<td>2.10 (1)</td>
<td>2.10 (0)</td>
</tr>
<tr>
<td>Grade 3, n (%)</td>
<td>2.0 (0.4)</td>
<td>3.0 (0.6)</td>
<td>2.0 (0.6)</td>
</tr>
<tr>
<td>Neutrophils, 10⁹/L</td>
<td>−1.0 (1.9)</td>
<td>−0.9 (2.0)</td>
<td>−1.2 (2.0)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>5.1 (1.1)</td>
<td>5.1 (1.0)†</td>
<td>1.0 (3.3)</td>
</tr>
<tr>
<td>Lymphocytes, 10⁹/L</td>
<td>−0.1 (0.6)</td>
<td>−0.1 (0.6)</td>
<td>0.3 (0.6)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>11 (2.3)*</td>
<td>6.1 (3.3)</td>
<td>2.0 (6.0)</td>
</tr>
<tr>
<td>Platelets, 10⁹/L</td>
<td>−30 (61.0)</td>
<td>−28 (62.4)</td>
<td>−34 (63.8)</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>6 (23.8)</td>
<td>4 (20.7)</td>
<td>6 (19.2)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>3.0 (0.6)</td>
<td>4.0 (0.8)</td>
<td>6.1 (3.3)</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>6.1 (16.8)</td>
<td>5 (14.0)</td>
<td>4.3 (12.2)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>3.0 (0.6)</td>
<td>2.0 (4.0)</td>
<td>2.0 (6.0)</td>
</tr>
<tr>
<td>Creatinine, mg/dL</td>
<td>0.1 (0.1)</td>
<td>0.1 (0.1)</td>
<td>0.0 (0.1)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>0.0 (1.2)</td>
<td>0.0 (2.4)</td>
<td>0.0 (1.0)</td>
</tr>
<tr>
<td>LDH, U/L</td>
<td>54 (89.5)</td>
<td>34 (64.4)</td>
<td>9.4 (70.1)</td>
</tr>
<tr>
<td>Grade 3 or 4, n (%)</td>
<td>4.0 (8.1)‡</td>
<td>2.0 (4.0)*</td>
<td>1.0 (3.3)</td>
</tr>
<tr>
<td>LDL cholesterol, mg/dL‡</td>
<td>15 (29.1)</td>
<td>12 (25.9)</td>
<td>7.2 (21.7)</td>
</tr>
<tr>
<td>% change</td>
<td>26 (53.3)</td>
<td>21 (23.9)</td>
<td>12 (22.6)</td>
</tr>
<tr>
<td>HDL cholesterol, mg/dL‡</td>
<td>12 (24.9)</td>
<td>15 (27.7)</td>
<td>9.2 (25.5)</td>
</tr>
<tr>
<td>% change</td>
<td>8 (21.7)</td>
<td>10 (24.5)</td>
<td>12 (26.8)</td>
</tr>
<tr>
<td>LDL:HDL ratio</td>
<td>−0.6 (31.1)</td>
<td>0.1 (36.4)</td>
<td>4.5 (23.6)</td>
</tr>
<tr>
<td>% change</td>
<td>−6.5 (23.0)</td>
<td>2.6 (23.6)</td>
<td>−6.5 (23.0)</td>
</tr>
</tbody>
</table>

Absolute values are presented as mean (SD) change from baseline at weeks 24 and 52 unless otherwise specified. Severity was graded using Common Terminology Criteria for Adverse Events Version 4.03.

*Grade 4 in one patient.
†Fasting values; not available for all patients.
‡Grade 4 in two patients.
§Fasting values; not available for all patients.

ADA, adalimumab; ALT, alanine aminotransferase; AST, aspartate aminotransferase; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; HDL, high-density lipoprotein; LDL, low-density lipoprotein; NA, not assessed; PBO, placebo.

Limitations

The study excluded patients with prior bDMARD failure, so data cannot be extrapolated to bDMARD-experienced patients; filgotinib was previously compared with placebo in this population. Generalisability to patients with less active RA is potentially limited because the study enrolled patients with moderate-to-severe disease. Placebo treatment was limited to 24 weeks due to ethical concerns. An elevated placebo response was observed, consistent with RA trial data showing increasing placebo rates over the last 20 years.34 In the present study, placebo response rates were especially high in geographical group B (predominantly Eastern Europe) and group C (Mexico and Argentina); as these groups comprised 65% of randomised patients, the regional differences contributed substantially to the overall placebo response rate. Nearly 50% of placebo-treated patients achieving study endpoints present a challenge to differentiating active agents from placebo. The study was not powered to compare AEs between arms, so no definitive conclusions about safety can be reached. Additional safety data will come from the integrated safety analysis across all phase II and III filgotinib trials, long-term extension study (ClinicalTrials.gov NCT03025308) and future registries.

CONCLUSIONS

In MTX-IR patients with active RA, filgotinib plus MTX reduced RA signs and symptoms, improved physical function and inhibited progression of structural joint damage. This study demonstrated non-inferiority of FIL200 plus MTX, but not FIL100 plus MTX, to adalimumab plus MTX, based on DAS28(CRP) low disease activity. Overall, filgotinib showed a favourable benefit-to-risk profile and both doses were well tolerated.

Author affiliations

1Rheumatology, CHU Montpellier, Montpellier, France
2Altoona Center for Clinical Research, Duncansville, Pennsylvania, USA
3First Department of Internal Medicine, School of Medicine, University of Occupational and Environmental Health, Kitakyushu, Japan
4Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
5Köhler & Milstein Research, Mérida, Mexico
6The Center for Rheumatology and Bone Research, Wheaton, Maryland, USA
7Rheumatology, All India Institute of Medical Sciences, New Delhi, India
8Gilead Sciences, Foster City, California, USA
9Galapagos NV, Mechelen, Belgium
10Medicine, Duke University Medical Center, Durham, North Carolina, USA
11Ichnos Sciences, New York, New York, USA
12Rheumatology and Clinical Immunology, Amsterdam University Medical Center, Amsterdam, The Netherlands
13Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Republic of Korea
14Medicine, University of Toronto, Mount Sinai Hospital, Toronto, Ontario, Canada
15Griffith University, Brisbane, Queensland, Australia

Acknowledgements We thank all patients and their families and the FINCH 1 investigators/site staff. Medical writing assistance was provided by Judith M Phillips, DVM, PhD, of Alphascientia, and Emmett Glass, PhD, MBA of Red Sky Medical, both funded by Gilead Sciences.
Contributors  FM, NM, JSS, CT and MCG contributed to the study concept and design. FM and MCG contributed to drafting the manuscript. YG and LY performed the statistical analyses. NM and JSS obtained funding. BC, AK, YT, DvdH, JAS, HSBB, UK, FM, BB, LB, YG, CT, JSS, AJ, MCg, NM, RBML, S-CB, EK and PN contributed to acquisition, analysis or interpretation of data; had full access to the data; reviewed the manuscript critically for important intellectual content; and approved the final version for publication.

Funding  The study was funded by Gilead Sciences.

Competing interests  BC received honoraria from AbbVie, Bristol-Myers Squibb, Gilead Sciences, Janssen, Lilly, Merck, Novartis, Pfizer, Roche-Chugai, Sanofi and UCB; and research grants from Novartis, Pfizer and Roche. AK received honoraria or consulting fees from AbbVie, Boehringer Ingelheim, Flexion, Genzyme, Gilead Sciences, Janssen, Novartis, Pfizer, Regeneron, Sanofi and Sun Pharma Advanced Research; was a paid instructor or speaker for AbbVie, Celgene, Flexion, Genzyme, Horizon, Merck, Novartis, Pfizer, Regeneron and Sanofi; and holds shares in Amgen, Gilead Sciences, GlaxoSmithKline, Pfizer and Sanofi. YT has received speaking fees and honoraria from AbbVie, Asahi Kasei, Astellas, Bristol-Myers, Chugai, Daiichi Sankyo, Eisai, Eli Lilly, Gilead, GSK, Janssen, Mitsubishi Tanabe, Novartis, Pfizer, Sanofi and YL Biologics; and research grants from AbbVie, Chugai, Daiichi Sankyo, Eisai, Mitsubishi Tanabe, Takeda and UCB. DvdH received consulting fees from AbbVie, Amgen, Astellas, AstraZeneca, BMS, Boehringer Ingelheim, Celgene, Cyzone, Daiichi Sankyo, Eisai, Eli Lilly, Galapagos, Gilead, GlaxoSmithKline, Janssen, Merck, Novartis, Pfizer, Regeneron, Roche, Sanofi, Takeda and UCB Pharma; and is the Director of Imaging Rheumatology BV. JAS, UK and S-CB report nothing to disclose. HSBB has received honoraria or consulting fees from AbbVie, Gilead Sciences, Horizon, Merck and research grants or support from AbbVie, Sanofi, Regeneron, Eli Lilly, Pfizer, Selecta Biosciences, Gilead Sciences, Horizon, Janssen and Pfizer. FM, BB, LB, YG and LY are employees and shareholders of Gilead Sciences. MCG has received honoraria or consulting fees from AbbVie, Amgen, Be/Gene, Genentech, Gilead Sciences, Lilly Pharmaceuticals, Sanofi Genzyme, RPharm and SetPoint. He is also an employee and shareholder of Galapagos NV. JSS, AJ and NM are former employees of Gilead Sciences and may hold shares. RBML has received honoraria or consulting fees from AbbVie, AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Galapagos NV, Novartis, Pfizer and UCB. EK has received honoraria or consulting fees from AbbVie, Amgen, AstraZeneca, Bristol-Myers Squibb, Celltrion, F Hoffmann-La Roche, Genentech, Gilead Sciences, Janssen, Lilly Pharmaceuticals, Merck, Myriad Immunomune, Pfizer, Sandoz, Sanofi Genzyme and Samsung Bioplocis has received speaking fees from AbbVie, Amgen, Bristol-Myers Squibb, F Hoffmann-La Roche, Janssen, Merck, Pfizer, Sanofi Genzyme and UCB; and has received research grants or support from AbbVie, Amgen, Gilead Sciences, Lilly Pharmaceuticals, Merck, Pfizer, PumaPharm and Sanofi. PN has received honoraria or consulting fees, grants or research support, or been a member of a speakers bureau for AbbVie, Bristol-Myers Squibb, Celgene, Gilead Sciences, Janssen, Lilly, MSD, Novartis, Pfizer, Roche, Sanofi and UCB.

Patient consent for publication  Not required.

Ethics approval  The trial was conducted in accordance with the Declaration of Helsinki and the International Council for Harmonisation guidelines. The protocol was approved by the institutional review board or ethics committee at each site.

Provenance and peer review  Not commissioned; externally peer reviewed.

Data availability statement  Data are available upon reasonable request. Anonymised individual patient data will be shared upon request for research purposes dependent upon the nature of the request, the merit of the proposed research, the availability of the data and the intended use. The full data sharing policy for Gilead Sciences can be found at https://www.gilead.com/about/ethics-and-code-of-conduct/policies.

Supplemental material  This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the context includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs  Bernard Combe http://orcid.org/0000-0003-4002-1861
Yoshiya Tanaka http://orcid.org/0000-0002-0807-7139

24 Westhovens R, Taylor PC, Alten R, et al. Filgotinib (GLPG0634/GS-6034), an oral JAK1 selective inhibitor, is effective in combination with methotrexate (MTX) in patients with...
Rheumatoid arthritis


# CLINICAL STUDY PROTOCOL

**Study Title:** A Randomized, Double-blind, Placebo- and Active-controlled, Multicenter, Phase 3 Study to Assess the Efficacy and Safety of Filgotinib Administered for 52 weeks in Combination with Methotrexate to Subjects with Moderately to Severely Active Rheumatoid Arthritis Who Have an Inadequate Response to Methotrexate

**Sponsors:** Gilead Sciences, Inc.  
333 Lakeside Drive  
Foster City, CA 94404  
USA

**IND Number:** 115,510  
**EudraCT Number:** 2016-000568-41  
**Clinical Trials.gov Identifier:** TBD  
**Indication:** Rheumatoid Arthritis  
**Protocol ID:** GS-US-417-0301  
**Gilead Clinical Program Manager Name:** Rebecca Pyle  
**Gilead Clinical Program Manager Telephone:** (650) 358-3630  
**Gilead Medical Monitor Name:** Franziska Matzkies  
**Gilead Medical Monitor Telephone:** (650) 389-8703  
**Gilead Medical Monitor Fax:** (650) 524-9297  
**Gilead Medical Monitor Mobile:** (310) 748-0404  
**Protocol Version/Date:** Original: 22 April 2016  
Amendment 1: 05 July 2016

## CONFIDENTIALITY STATEMENT

The information contained in this document, particularly unpublished data, is the property or under control of Gilead Sciences, Inc., and is provided to you in confidence as an investigator, potential investigator, or consultant, for review by you, your staff, and an applicable Institutional Review Board or Independent Ethics Committee. The information is only to be used by you in connection with authorized clinical studies of the investigational drug described in the protocol. You will not disclose any of the information to others without written authorization from Gilead Sciences, Inc., except to the extent necessary to obtain informed consent from those persons to whom the drug may be administered.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>2</td>
</tr>
<tr>
<td>LIST OF IN-TEXT TABLES</td>
<td>5</td>
</tr>
<tr>
<td>LIST OF IN-TEXT FIGURES</td>
<td>5</td>
</tr>
<tr>
<td>PROTOCOL SYNOPSIS</td>
<td>6</td>
</tr>
<tr>
<td>GLOSSARY OF ABBREVIATIONS AND DEFINITION OF TERMS</td>
<td>16</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>19</td>
</tr>
<tr>
<td>1.1. Background</td>
<td>19</td>
</tr>
<tr>
<td>1.2. Filgotinib</td>
<td>20</td>
</tr>
<tr>
<td>1.2.1. General Information</td>
<td>20</td>
</tr>
<tr>
<td>1.2.2. Nonclinical Pharmacology, Absorption, Distribution, Metabolism, and Excretion (ADME) and Toxicology</td>
<td>20</td>
</tr>
<tr>
<td>1.2.3. Clinical Trials of Filgotinib</td>
<td>22</td>
</tr>
<tr>
<td>1.3. Information about Humira® (adalimumab)</td>
<td>27</td>
</tr>
<tr>
<td>1.4. Rationale for this Study</td>
<td>27</td>
</tr>
<tr>
<td>1.4.1. Rationale for Study Design</td>
<td>28</td>
</tr>
<tr>
<td>1.4.2. Rationale for the Outcome Measures</td>
<td>29</td>
</tr>
<tr>
<td>1.4.3. Rationale for the Choice of Dose and Dosing Interval</td>
<td>30</td>
</tr>
<tr>
<td>1.5. Risk/Benefit Assessment for the Study</td>
<td>31</td>
</tr>
<tr>
<td>1.6. Compliance</td>
<td>32</td>
</tr>
<tr>
<td>2. OBJECTIVES</td>
<td>33</td>
</tr>
<tr>
<td>3. STUDY DESIGN</td>
<td>34</td>
</tr>
<tr>
<td>3.1. Endpoints</td>
<td>34</td>
</tr>
<tr>
<td>3.2. Study Design</td>
<td>35</td>
</tr>
<tr>
<td>3.3. Study Treatments</td>
<td>37</td>
</tr>
<tr>
<td>3.4. Duration of Treatment</td>
<td>38</td>
</tr>
<tr>
<td>3.5. Criteria for Interruption or Discontinuation of Study Treatment</td>
<td>38</td>
</tr>
<tr>
<td>3.5.1. Study drug interruption considerations</td>
<td>38</td>
</tr>
<tr>
<td>3.5.2. Study drug discontinuation considerations</td>
<td>38</td>
</tr>
<tr>
<td>3.6. End of Study</td>
<td>40</td>
</tr>
<tr>
<td>3.7. Post Study Care</td>
<td>40</td>
</tr>
<tr>
<td>3.8. Biomarker Samples</td>
<td>40</td>
</tr>
<tr>
<td>3.8.1. Biomarker Samples to Address the Study objectives</td>
<td>40</td>
</tr>
<tr>
<td>3.8.2. Biomarker Samples for Optional Future Research</td>
<td>41</td>
</tr>
<tr>
<td>3.8.3. Biomarker Samples for Optional Genomic Research</td>
<td>42</td>
</tr>
<tr>
<td>3.9. SUBSTUDY: Carotid Artery Ultrasound</td>
<td>43</td>
</tr>
<tr>
<td>3.10. SUBSTUDY: MRI of Hand/Wrist</td>
<td>43</td>
</tr>
<tr>
<td>4. SUBJECT POPULATION</td>
<td>44</td>
</tr>
<tr>
<td>4.1. Number of Subjects and Subject Selection</td>
<td>44</td>
</tr>
<tr>
<td>4.2. Inclusion Criteria</td>
<td>44</td>
</tr>
<tr>
<td>4.3. Exclusion Criteria</td>
<td>46</td>
</tr>
<tr>
<td>4.4. MRI Sub-Study Criteria</td>
<td>49</td>
</tr>
<tr>
<td>4.4.1. Inclusion criteria for MRI subsudy</td>
<td>49</td>
</tr>
<tr>
<td>4.4.2. Exclusion criteria for MRI subsudy</td>
<td>50</td>
</tr>
<tr>
<td>5. INVESTIGATIONAL MEDICINAL PRODUCTS</td>
<td>51</td>
</tr>
<tr>
<td>5.1. Randomization, Blinding and Treatment Codes</td>
<td>51</td>
</tr>
</tbody>
</table>
5.1. Procedures for Breaking Treatment Codes .......................................................... 51
5.2. Description and Handling of Filgotinib and PTM Filgotinib .................................... 51
5.2.1. Formulation of Filgotinib and PTM Filgotinib ..................................................... 51
5.2.2. Packaging and Labeling of Filgotinib and PTM Filgotinib ............................... 52
5.2.3. Storage and Handling of Filgotinib and PTM Filgotinib ................................. 52
5.2.4. Dosage and Administration of Filgotinib ......................................................... 52
5.3. Description and Handling of Adalimumab and PTM Adalimumab ......................... 53
5.3.1. Formulation of Adalimumab and PTM Adalimumab ........................................ 53
5.3.2. Packaging and Labeling of Adalimumab and PTM Adalimumab ....................... 53
5.3.3. Storage and Handling of Adalimumab and PTM Adalimumab ......................... 53
5.3.4. Dosage and Administration of Adalimumab ...................................................... 53
5.4. Description and Handling of Methotrexate (MTX) .............................................. 54
5.4.1. Formulation of MTX ......................................................................................... 54
5.4.2. Packaging and Labeling of MTX ..................................................................... 54
5.4.3. Storage and Handling of MTX ......................................................................... 54
5.4.4. Dosage and Administration of MTX .................................................................. 54
5.5. Prior and Concomitant Medications ...................................................................... 54
5.6. Accountability for Study Drugs ............................................................................ 57
5.6.1. Investigational Medicinal Product Return or Disposal ...................................... 58

6. STUDY PROCEDURES ......................................................................................... 59
6.1. Subject Enrollment and Treatment Assignment .................................................... 59
6.2. Pretreatment Assessments ................................................................................... 59
6.2.1. Screening Visit .................................................................................................. 59
6.2.2. Day 1 Assessments ......................................................................................... 61
6.3. Randomization ..................................................................................................... 62
6.4. Week 2 through Week 44 Assessments ................................................................ 62
6.5. Week 52/Early Termination (ET) ........................................................................ 64
6.6. Post Treatment Week 4 ....................................................................................... 66
6.7. Study Assessments ............................................................................................... 66
6.7.1. Efficacy ........................................................................................................... 66
6.7.2. Safety and Tolerability .................................................................................... 70
6.7.3. Treatment Satisfaction Questionnaire for Medication (TSQM) ...................... 70
6.7.4. Clinical Laboratory Evaluations ...................................................................... 70
6.7.5. Vital Signs ....................................................................................................... 71
6.7.6. Physical Examination ..................................................................................... 71
6.7.7. Other Safety Assessments .............................................................................. 71
6.8. Pharmacokinetics Assessments .......................................................................... 72
6.9. Biomarker Assessments ...................................................................................... 72

7. ADVERSE EVENTS AND TOXICITY MANAGEMENT ...................................... 73
7.1. Definitions of Adverse Events, Adverse Reactions, and Serious Adverse Events .... 73
7.1.1. Adverse Events ................................................................................................ 73
7.1.2. Serious Adverse Events .................................................................................. 73
7.1.3. Clinical Laboratory Abnormalities and Other Abnormal Assessments as Adverse Events or Serious Adverse Events .......................................................... 74
7.2. Assessment of Adverse Events and Serious Adverse Events ................................ 74
7.2.1. Assessment of Causality for Study Drugs and Procedures ............................. 74
7.2.2. Assessment of Severity .................................................................................. 75
7.3. Investigator Requirements and Instructions for Reporting Adverse Events and Serious Adverse Events to Gilead ........................................................................ 75
7.4. Gilead Reporting Requirements ........................................................................... 77
7.5. Clinical Laboratory Abnormalities and Other Abnormal Assessments as Adverse Events or Serious Adverse Events .......................................................... 77

CONFIDENTIAL Page 3 05 July 2016

Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Appendix 1. Investigator Signature Page ........................................................................................................100
Appendix 2. Study Procedures Table ..................................................................................................................101
Appendix 3. Management of Clinical and Laboratory Adverse Events ..............................................................106
Appendix 4. Common Terminology Criteria for Adverse Events (CTCAE) v4.03 ................................................107
Appendix 5. Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Methods ........................................................................................................................108
Appendix 6. Laboratory Assessment Table ........................................................................................................112
Appendix 7. List of Joints to be Evaluated (66/68 Joint Count) ........................................................................113
Appendix 8. The 2010 American College of Rheumatology-European League Against Rheumatism Collaborative Initiative Classification Criteria for Rheumatoid Arthritis (Aletaha et al 2010) .........................................................................................................................114
Appendix 9. American College of Rheumatology Response Evaluations/Preliminary Definition of Improvement in Rheumatoid Arthritis (Felson et al 1995) ..................................................................................................................116
Appendix 10. Disease Activity Score (DAS28) {Prevoo et al 1995} ......................................................................118
Appendix 11. Procedures and Specifications ......................................................................................................119

LIST OF IN-TEXT TABLES

Table 1-1. Summary and analysis of ACR20/50/70 response at Weeks 12 and 24 (NRI [ITT Population]); GLPG0634-CL-203 ................................................................................................................23
Table 1-2. Summary and analysis of ACR20/50/70 response at Weeks 12 and 24 (NRI [ITT Population]); GLPG0634-0204 ................................................................................................................26
Table 7-1. Grading of Adverse Event Severity ....................................................................................................75

LIST OF IN-TEXT FIGURES

Figure 3-1. Study Design ................................................................................................................................37
**Study Title:** A Randomized, Double-blind, Placebo- and Active-controlled, Multicenter, Phase 3 Study to Assess the Efficacy and Safety of Filgotinib Administered for 52 weeks in Combination with Methotrexate to Subjects with Moderately to Severely Active Rheumatoid Arthritis who have an Inadequate Response to Methotrexate

**IND Number:** 115,510
**EudraCT Number:** 2016-000568-41
**Clinical Trials.gov Identifier:** TBD

**Study Centers Planned:** Approximately 250 - 300 centers worldwide

**Objectives:**
- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of rheumatoid arthritis (RA) as measured by the proportion of subjects achieving an American College of Rheumatology 20% improvement response (ACR20) at Week 12.

The secondary objectives of this study include:
- To evaluate the effects of filgotinib versus placebo as measured by the proportion of subjects achieving Disease Activity Score for 28 joint count using c-reactive protein (DAS28(CRP)) ≥ 2.6 at Week 12
- To evaluate the effect of filgotinib versus placebo on physical function as measured by change from Baseline in the Health Assessment Questionnaire Disability Index (HAQ-DI) score at Week 12
- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP)<2.6 at Week 24
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

To evaluate the effects of filgotinib versus placebo on preservation of joint structure as measured by change from Baseline in the van der Heijde modified Total Sharp Score (mTSS) at Week 24
To evaluate the effects of filgotinib versus adalimumab for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP) ≤3.2 at Week 12
To evaluate the safety and tolerability of filgotinib
To evaluate the effects of filgotinib on work productivity, fatigue, and general quality of life as measured by SF-36, FACIT-Fatigue, EQ-5D and WPAI-RA

The exploratory objectives of this study include:
To characterize the pharmacokinetics (PK) of filgotinib and its metabolite (GS-829845, formerly G254445)
To characterize the association of host genetics and other markers with disease severity, disease progression and treatment response to filgotinib in subjects with RA
To evaluate the effects of filgotinib on healthcare resource utilization and other patient reported outcomes

Study Design:
This is a randomized, double-blind, placebo- and active-controlled, Phase 3 study in adult male and female subjects with active RA who have an inadequate response to MTX (MTX-IR). The study is designed to evaluate the efficacy, safety, and tolerability of filgotinib as well as its effect on work productivity, fatigue, and quality of life.
Approximately 1650 subjects will be randomized in a 3:3:2:3 ratio to filgotinib 200 mg, filgotinib 100 mg, active comparator (adalimumab), or placebo to match (PTM) administered for up to 52 weeks, all in the context of a weekly stable dose of MTX:
Filgotinib 200 mg group: filgotinib (200 mg q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
Filgotinib 100 mg group: filgotinib (100 mg q.d.) + PTM filgotinib 200 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
Active comparator group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + adalimumab (40 mg s.c. injection q2w) (N=300)
Placebo control group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
Randomization will be stratified by geographic region, prior exposure to biologic disease modifying antirheumatic drugs (bDMARDs), and presence of rheumatoid factor (RF) or anti-CCP (cyclic citrullinated peptide) antibody (Ab) at screening.

At Week 14, subjects who have not achieved at least 20% improvement from Day 1 in both swollen joint count (SJ C) and tender joint count (TJC) will discontinue investigational study drug dosing but will continue with study visits and assessments per protocol. All subjects meeting this criterion who discontinue from investigational therapy are to receive standard of care treatment for their RA as determined by the investigator.

At Week 24, all subjects assigned to placebo + MTX will be reassigned 1:1 to either filgotinib 100 mg q.d. or 200 mg q.d. in addition to MTX in a blinded fashion and will continue in the study per protocol up to Week 52.

All subjects who continue on study drug will be evaluated for loss of therapeutic response from Week 30 through Week 52. Subjects failing to maintain at least a 20% improvement from Day 1 in TJC and SJ C, (which is confirmed at 2 consecutive visits), will discontinue from investigational study drug therapy but will continue with study visits and assessments per protocol. All subjects meeting this criterion who discontinue from investigational study drug dosing are to receive standard of care treatment for their RA as determined by the investigator.

All subjects who have received at least one dose of study drug and exit the study early will complete an early termination (ET) visit at the time of study discontinuation, with a follow up visit four weeks after the last dose of study drug (Post Treatment Week 4), regardless of dosing duration.

At completion of the 52-week dosing period, subjects who have not discontinued assigned study drug dosing, will be provided the option to enroll into a separate Long Term Extension (LTE) study (GS-US-417-0304).
Study Design

Number of Subjects Planned: Approximately 1650 subjects

Target Population: Male or female subjects who are ≥18 years of age with a diagnosis of RA meeting the 2010 ACR/European League Against Rheumatism (EULAR) criteria, who have active RA despite treatment with MTX

Duration of Treatment: Up to 52 weeks

Diagnosis and Main Eligibility Criteria: For a complete list of study inclusion and exclusion criteria, please refer to Sections 4.2 and 4.3.

Main Eligibility Criteria

1) Male or female subjects who are ≥18 years of age, on the day of signing informed consent.

2) Have a diagnosis of RA (2010 ACR/EULAR criteria), and are ACR functional class I-III.

3) Have ≥6 swollen joints (from a SJ C66) and ≥6 tender joints (from a TJ C68) at both Screening and Day 1.

4) Must meet at least one of the following parameters at Screening:

   a) ≥1 documented joint erosion on radiographs of the hands, wrists or feet by central reading AND a positive result for anti-CCP or RF at the central laboratory,

   OR
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

b) 3 documented erosions on radiographs of the hands, wrists or feet by central reading if both antibodies (ie, RF, anti-CCP) are negative (based on central laboratory),
OR

c) Serum CRP > 6 mg/L based on central laboratory value

5) Ongoing treatment with a stable dose of MTX as described below:

a) Use of oral MTX on a continuous basis for at least 12 weeks prior to Day 1 and on a stably prescribed dose of 7.5-25 mg/weekly for at least 4 weeks prior to Day 1. Stable doses of <7.5 mg/week are allowed only in the presence of intolerance or toxicity to higher doses or where higher doses are prohibited by the local label or local clinical practice. Doses >25 mg weekly are not permitted during the study.

b) Subjects should be receiving an adequate and prescribed stable dose of folic acid (5 mg/week total dose or as per local clinical practice) which should be confirmed or initiated at Screening, and continued throughout the study.

c) Subjects may use concomitant hydroxychloroquine (HCQ) ≥400 mg/day or chloroquine ≥250 mg/day during the study with the prescription having been stable for at least 4 weeks prior to Day 1.

6) Subjects that have failed prior therapy with a bDMARD are not eligible to participate. Subjects with prior exposure to one bDMARD may be enrolled (approximately 20% of total study population) if there is documented evidence of limited exposure (ie, less than 3 months) to the bDMARD.

Study Procedures/ Frequency:

Participating subjects will visit the clinical study center at Screening, Day 1, Weeks 2, 4, 8, 12, 14, 16, 20, 24, 26, 30, 36, 44, and 52 or ET. For those subjects not entering the LTE, a follow-up visit will be planned 4 weeks after the last dose of study drug (Post Treatment Week 4). Consequently, subjects are planned to be in the study for approximately 60 weeks.

During the Screening period, the subject’s radiographs (hands, wrists, and feet) will be sent for blinded central scoring. In addition, information on demographics, medical history/concomitant diseases, prior and current RA medication will be collected. A physical examination, a 12-lead electrocardiogram (ECG), and clinical laboratory assessments will be conducted to determine the subject’s eligibility for study participation. The screening window may be extended to up to 42 days prior to the Day 1 visit for subjects who require repeat collection of radiographs only.
At Day 1, after the subject’s eligibility for the study has been confirmed, the subject will be randomized into the study to receive filgotinib 200 mg q.d., filgotinib 100 mg q.d., active comparator (adalimumab 40 mg q2w), or placebo control. All subjects will continue their stable dose of MTX weekly. Study medication will be dispensed and subjects and/or their care-giver will be instructed on administration of s.c. injections of adalimumab/PTM.

On-treatment assessments will include: adverse events (AEs), concomitant medications, review of study medication compliance through drug accountability, SJC66, TJC68, Subject’s Global Assessment, Subject’s Pain Assessment, HAQ-DI, Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue), Short-Form Health Survey (SF-36), Work Productivity and Activity Impairment- Rheumatoid Arthritis (WPAI-RA), EuroQol 5 Dimensions (EQ-5D), Treatment Satisfaction Questionnaire for Medication (TSQM), Exploratory Patient Reported Outcomes, Healthcare Resource Utilization, Physician’s Global Assessment, physical examination, weight, vital signs, serum CRP, blood and urine sampling for safety laboratory tests and biomarkers at selected visits, and urine pregnancy tests (females of child bearing potential only).

A resting 12-lead ECG should be performed at Screening, Weeks 12, 24, 36, and Week 52 or ET, and at the Post Dosing Week 4 follow-up visit, as applicable. ECGs should be interpreted by the investigator (or qualified designee) for clinical significance.

Blood samples for PK analysis should be collected at least 30 minutes post dose on Week 4, prior to study drug administration on Weeks 12 and 24, and anytime at the final assessment visit on Week 52 or ET.

For subjects who consent to participate in the optional PK substudy (approximately 60 subjects), the daily dose of study drug should be administered under supervision in the clinic (between Weeks 2 and 8, inclusive), and additional PK samples should be collected predose and at 0.5, 1, 2, 3, 4, and 6 hours post dose.

Blood samples for biomarker analysis should be collected at Day 1, Weeks 4, 8, 12, 24, 36, and Week 52 or ET and urine samples should be collected at Day 1, Weeks 12, 24, and Week 52 or ET.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Radiographs of the bilateral hands, wrists, and feet should be taken at Screening, Weeks 12, 24, and Week 52 or ET. However, if a subject prematurely discontinues study participation and has had X-rays performed within the previous 12 weeks, then X-rays do not need to be repeated at the ET visit. Radiographs after enrollment may be done +/- 7 days of the scheduled visit. Post treatment follow-up assessments include AEs, concomitant medications, physical examination, 12-lead ECG, weight, vital signs, blood and urine sampling for safety laboratory tests, and urine pregnancy tests (females of child bearing potential only).

Optional MRI sub-study
At selected sites, a subset of subjects may consent to participate in an optional MRI sub-study (see Section 3.9). For those subjects, an MRI of a single hand and wrist will be performed post-randomization within 7 days of the first dose and repeated at Week 12. Hand and wrist of each subject will be imaged using a 1.5 Tesla or 3 Tesla whole-body MRI scanner. Images will be centrally read by 2 radiologists who are blinded with regard to visit sequence, subject treatment assignment and the alternate reader’s scores.

Test Product, Dose, and Mode of Administration:
200 mg filgotinib orally q.d.
100 mg filgotinib orally q.d.

Reference Therapy, Dose, and Mode of Administration:
Adalimumab injection (40 mg s.c.) or PTM s.c. q2w
PTM filgotinib 200 mg orally q.d.
PTM filgotinib 100 mg orally q.d.

Required Background Medication:
Stable prescription of weekly MTX as described in the inclusion criteria and administered according to the local product label/standard of care.

Criteria for Evaluation:
Safety:
Safety will be assessed by documentation of AEs, clinical laboratory tests, physical examinations, vital signs, and 12-lead ECGs during the study.
Efficacy: The primary endpoint is the proportion of subjects who achieve an ACR20 response at Week 12.

The key secondary endpoints are:

- The proportion of subjects who achieve DAS28 (CRP)\(\geq 3.2\) at Week 12
- Change from Baseline in the HAQ-DI score at Week 12
- The proportion of subjects who achieve DAS28 (CRP)<2.6 at Week 24
- Change from Baseline in mTSS at Week 24

Pharmacokinetics: Plasma concentrations of filgotinib and its metabolite (GS-829845) will be analyzed.

Statistical Methods: The primary analysis set for efficacy analyses will be the Full Analysis Set (FAS), which includes all randomized subjects who received at least one dose of study drug.

The primary endpoint is the proportion of subjects who achieve an ACR20 response at Week 12. The primary analysis will consist of a superiority test of filgotinib 200 mg compared to placebo based on the ACR20 response rate at Week 12. Cochran-Mantel-Haenszel approach adjusting for the randomization stratification factors will be used for the hypothesis testing at the 2-sided 0.05-level.

The following hypothesis testing will commence after the primary analysis reaches statistical significance, and will be tested according to the hierarchical testing principle at the 2-sided 0.05-level. If a null hypothesis is not rejected, formal sequential testing will be stopped and only nominal significance will be reported for the remaining hypotheses:

1) Superiority of filgotinib 200 mg compared to placebo based on the response rate of DAS28 (CRP)\(\geq 3.2\) at Week 12
2) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in HAQ-DI at Week 12
3) Superiority of filgotinib 200 mg compared to placebo based on the response rate of DAS28 (CRP)<2.6 at Week 24
4) Superiority of filgotinib 100 mg compared to placebo based on the ACR20 response rate at Week 12
5) Superiority of filgotinib 100 mg compared to placebo based on the response rate of DAS28 (CRP)\(\geq 3.2\) at Week 12
6) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in HAQ-DI at Week 12

7) Superiority of filgotinib 100 mg compared to placebo based on the response rate of DAS28 (CRP)<2.6 at Week 24

8) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in mTSS at Week 24

9) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in mTSS at Week 24

10) Non-inferiority of filgotinib 200 mg compared to adalimumab based on the response rate of DAS28 (CRP)âB.2 at Week 12

11) Non-inferiority of filgotinib 100 mg compared to adalimumab based on the response rate of DAS28 (CRP)âB.2 at Week 12

If non-inferiority of at least one filgotinib dose compared to adalimumab based on the response rate of DAS28 (CRP)âB.2 at Week 12 is demonstrated, then a superiority test of the respective filgotinib dose compared to adalimumab will be performed.

All continuous endpoints will be summarized using an 8-number summary (n, mean, standard deviation [SD], median, 1st quartile [Q1], 3rd quartile [Q3], minimum, maximum) by treatment group. All categorical endpoints will be summarized by the number and percentage of subjects who meet the endpoint definition.

Safety endpoints will be analyzed by the number and percent of subjects with events or abnormalities for categorical values or 8-number summary (n, mean, SD, median, Q1, Q3, minimum, maximum) for continuous data by treatment group.

Sample size is determined based on the superiority test of filgotinib 200 mg compared to placebo on the change from Baseline in mTSS at Week 24. When assuming a difference of 0.4 between filgotinib and placebo on change from Baseline in mTSS at Week 24 and a common standard deviation of 1.85, 450 subjects in each of the filgotinib group and placebo control group are required to obtain 90% power at a 2-sided 0.05-level.

A sample size of 450 subjects in each of the filgotinib groups and placebo control group will provide over 95% power to detect an increase in ACR20 response rate from 45% to 65% between the placebo control group and the filgotinib group respectively, using a 2-sided 0.05-level test.
Based on {Liu et al 2014}, 450 subjects in each of the filgotinib 200 mg and placebo group, and 300 subjects in the adalimumab group, will provide over 90% power at 2-sided 0.05 significance level to demonstrate that filgotinib 200 mg preserves more than 50% of the effect of adalimumab with respect to the response rate of DAS28 (CRP)\(\times B.2\) at Week 12, assuming both filgotinib 200 mg and adalimumab groups have similar response rates of DAS28 (CRP)\(\times B.2\). Given this study has a placebo group, assay sensitivity can be established through a direct comparison of adalimumab to placebo.

In summary, the total sample size will be 1650 (450 subjects per group for filgotinib 100 mg, 200 mg and placebo groups; 300 subjects for active comparator group).

This study will be conducted in accordance with the guidelines of Good Clinical Practice (GCP) including archiving of essential documents.
**GLOSSARY OF ABBREVIATIONS AND DEFINITION OF TERMS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab</td>
<td>Antibody</td>
</tr>
<tr>
<td>ACR X</td>
<td>American College of Rheumatology X% improvement</td>
</tr>
<tr>
<td>AE</td>
<td>adverse event</td>
</tr>
<tr>
<td>ALT</td>
<td>alanine aminotransferase</td>
</tr>
<tr>
<td>ANA</td>
<td>anti-nuclear antibody</td>
</tr>
<tr>
<td>aPTT</td>
<td>activated partial thromboplastin time</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>AST</td>
<td>aspartate aminotransferase</td>
</tr>
<tr>
<td>AUC0-12h</td>
<td>area under the plasma drug concentration-time curve of a dosing interval</td>
</tr>
<tr>
<td>bDMARD(s)</td>
<td>biologic disease modifying antirheumatic drug(s)</td>
</tr>
<tr>
<td>b.i.d.</td>
<td><em>bis in diem</em>, twice daily</td>
</tr>
<tr>
<td>C12</td>
<td>trough plasma concentration (just before the next dosing ie predose sample)</td>
</tr>
<tr>
<td>CCP</td>
<td>cyclic citrullinated peptide</td>
</tr>
<tr>
<td>CD</td>
<td>Crohn's Disease</td>
</tr>
<tr>
<td>CDAI</td>
<td>Clinical Disease Activity Index</td>
</tr>
<tr>
<td>csDMARD(s)</td>
<td>Conventional synthetic disease modifying antirheumatic drug(s)</td>
</tr>
<tr>
<td>CES</td>
<td>Carboxylesterases</td>
</tr>
<tr>
<td>CIA</td>
<td>collagen-induced arthritis</td>
</tr>
<tr>
<td>Cmax</td>
<td>maximum observed plasma concentration</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>CRO</td>
<td>Contract Research Organization</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>CV-SEAC</td>
<td>Cardiovascular Safety Endpoint Adjudication Committee</td>
</tr>
<tr>
<td>CY P</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>DAS28</td>
<td>Disease Activity Score based on 28 joints</td>
</tr>
<tr>
<td>DBP</td>
<td>diastolic blood pressure</td>
</tr>
<tr>
<td>DMARD(s)</td>
<td>disease-modifying antirheumatic drugs</td>
</tr>
<tr>
<td>DMC</td>
<td>data monitoring committee</td>
</tr>
<tr>
<td>DSS</td>
<td>Dextran sulphate sodium</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>eCRF</td>
<td>electronic case report form</td>
</tr>
<tr>
<td>ET</td>
<td>early termination</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>EULAR</td>
<td>European League Against Rheumatism</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>EuroQol 5 Dimensions</td>
</tr>
<tr>
<td>FACIT-Fatigue</td>
<td>Functional Assessment of Chronic Illness Therapy-Fatigue</td>
</tr>
<tr>
<td>FAS</td>
<td>Full Analysis Set</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FSH</td>
<td>follicle stimulating hormone</td>
</tr>
</tbody>
</table>
Filgotinib  
Protocol GS-US-417-0301  
Gilead Sciences, Inc.  

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCP</td>
<td>Good Clinical Practice</td>
</tr>
<tr>
<td>GFR</td>
<td>Glomerular filtration rate</td>
</tr>
<tr>
<td>GGT</td>
<td>gamma glutamyl transferase</td>
</tr>
<tr>
<td>GI</td>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>GLP</td>
<td>Galapagos</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>Health Assessment Questionnaire - Disability Index</td>
</tr>
<tr>
<td>HCV</td>
<td>hepatitis C virus</td>
</tr>
<tr>
<td>HCG</td>
<td>human chorionic gonadotropin</td>
</tr>
<tr>
<td>HCQ</td>
<td>Hydroxychloroquine</td>
</tr>
<tr>
<td>HDL</td>
<td>high density lipoprotein</td>
</tr>
<tr>
<td>hERG</td>
<td>Human ether-a-gogo related gene</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HR</td>
<td>heart rate</td>
</tr>
<tr>
<td>IC_{50}</td>
<td>half maximal inhibitory concentration</td>
</tr>
<tr>
<td>ICF</td>
<td>informed consent form</td>
</tr>
<tr>
<td>ICH</td>
<td>International Council for Harmonization</td>
</tr>
<tr>
<td>IEC</td>
<td>Independent Ethics Committee</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>IMP</td>
<td>investigational medicinal product</td>
</tr>
<tr>
<td>INR</td>
<td>international normalized ratio</td>
</tr>
<tr>
<td>IRB</td>
<td>Independent Review Board</td>
</tr>
<tr>
<td>ITT</td>
<td>intent-to-treat</td>
</tr>
<tr>
<td>IV</td>
<td>Intravenous</td>
</tr>
<tr>
<td>IWRS</td>
<td>interactive web response system</td>
</tr>
<tr>
<td>JAK</td>
<td>janus kinase</td>
</tr>
<tr>
<td>LCMS/MS</td>
<td>liquid chromatography mass spectrometry</td>
</tr>
<tr>
<td>LDL</td>
<td>low density lipoprotein</td>
</tr>
<tr>
<td>LH</td>
<td>luteinizing hormone</td>
</tr>
<tr>
<td>LTE</td>
<td>long term extension</td>
</tr>
<tr>
<td>MAA</td>
<td>Marketing Authorization Application</td>
</tr>
<tr>
<td>MACE</td>
<td>major adverse cardiovascular event</td>
</tr>
<tr>
<td>MCV</td>
<td>mean corpuscular volume</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnet resonance imaging</td>
</tr>
<tr>
<td>MTX</td>
<td>Methotrexate</td>
</tr>
<tr>
<td>MTX-IR</td>
<td>inadequate response to Methotrexate</td>
</tr>
<tr>
<td>mTSS</td>
<td>modified Total Sharp Score</td>
</tr>
<tr>
<td>NDA</td>
<td>New Drug Application</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>nonsteroidal anti-inflammatory drugs</td>
</tr>
<tr>
<td>NYHA</td>
<td>New York Heart Association</td>
</tr>
<tr>
<td>OATs</td>
<td>Organic anion transporters</td>
</tr>
<tr>
<td>PBMC</td>
<td>Peripheral blood mononuclear cells</td>
</tr>
</tbody>
</table>

CONFIDENTIAL  
Page 17  
05 July 2016
Filgotinib Protocol GS-US-417-0301
Gilead Sciences, Inc.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>Pharmacodynamics</td>
</tr>
<tr>
<td>PK</td>
<td>Pharmacokinetics</td>
</tr>
<tr>
<td>PP</td>
<td>per-protocol</td>
</tr>
<tr>
<td>PT</td>
<td>prothrombin time</td>
</tr>
<tr>
<td>PTM</td>
<td>Placebo to match</td>
</tr>
<tr>
<td>q.d.</td>
<td><em>quaque die</em> once daily</td>
</tr>
<tr>
<td>RA</td>
<td>rheumatoid arthritis</td>
</tr>
<tr>
<td>RF</td>
<td>rheumatoid factor</td>
</tr>
<tr>
<td>RR</td>
<td>respiratory rate</td>
</tr>
<tr>
<td>SAE</td>
<td>serious adverse event</td>
</tr>
<tr>
<td>SAP</td>
<td>Statistical Analysis Plan</td>
</tr>
<tr>
<td>SBP</td>
<td>systolic blood pressure</td>
</tr>
<tr>
<td>s.c.</td>
<td>Subcutaneous</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SDAI</td>
<td>Simplified Diagnostic Activity Index</td>
</tr>
<tr>
<td>SF-36</td>
<td>Short-form health survey</td>
</tr>
<tr>
<td>SI</td>
<td>International system of units</td>
</tr>
<tr>
<td>SJC66</td>
<td>swollen joint count based on 66 joints</td>
</tr>
<tr>
<td>SOC</td>
<td>system organ class</td>
</tr>
<tr>
<td>STIR</td>
<td>short tau inversion recovery sequence</td>
</tr>
<tr>
<td>TB</td>
<td>Tuberculosis</td>
</tr>
<tr>
<td>TEAEs</td>
<td>Treatment emergent adverse events</td>
</tr>
<tr>
<td>TJC68</td>
<td>tender joint count based on 68 joints</td>
</tr>
<tr>
<td>t\text{max}</td>
<td>the time of occurrence of maximum observed plasma concentration</td>
</tr>
<tr>
<td>TNF</td>
<td>tumor necrosis factor alpha</td>
</tr>
<tr>
<td>TNFi</td>
<td>tumor necrosis factor inhibitors</td>
</tr>
<tr>
<td>UGT\text{i}s</td>
<td>uridine 5'-disphosphate glucuronosyltransferases</td>
</tr>
<tr>
<td>ULN</td>
<td>upper limit of normal</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>vPBMCS</td>
<td>Viably frozen peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>vs.</td>
<td>Versus</td>
</tr>
<tr>
<td>WBC</td>
<td>white blood cell</td>
</tr>
<tr>
<td>WPAI-RA</td>
<td>Work Productivity and Activity Impairment ^ Rheumatoid Arthritis</td>
</tr>
</tbody>
</table>

**Definition of Terms**

**QTcF** QT interval corrected for HR according to Fridericia’s formula: 

\[
\text{QTcF} = \frac{\text{QT}}{(\text{RR})^{(1/3)}}, \quad \text{where RR} = 60/\text{HR} \\
\text{RR} = \text{RR interval in seconds} \\
\text{HR} = \text{heart rate in beats per minute}
\]
1. INTRODUCTION

1.1. Background

Rheumatoid arthritis (RA) is a chronic, systemic inflammatory disease that affects approximately 1.3 million adults in the United States (US) \( \text{Helmick et al. 2008} \). Rheumatoid arthritis manifests principally as an attack on peripheral joints and may lead to marked destruction and deformity of joints, with considerable disability and impact on quality of life. It is characterized by the production of autoantibodies, synovial inflammation with formation of pannus tissue, and erosion of underlying cartilage and bone. Although people of any age can be affected, the onset of RA is most frequent between the ages of 40 and 50 years, and women are affected 3 times more often than men. While the cause of RA is still not completely understood, aberrant B-cell activation, T-cell co-stimulation, osteoclast differentiation, and cytokine release all have been implicated in its pathogenesis.

Treatment of RA is dependent on severity, the patient’s co-morbidities and initial response to therapy. Methotrexate (MTX) is a conventional disease modifying anti-rheumatic drug (DMARD) and continues to be the cornerstone of RA therapy \( \text{Singh et al. 2012} \). Patients with an inadequate response to conventional DMARD(s) are often treated with biologic therapies such as tumor necrosis factor inhibitors (TNFi) as an initial second line therapy. However, approximately 28% to 58% of RA patients with inadequate response to MTX fail TNFi as reviewed in \( \text{Redlich et al. 2003} \). In this setting, treatment guidelines recommend either switching to another TNFi, alternate biologic, or to a small molecule drug \( \text{Singh et al. 2012} \). Despite significant advances in disease management in recent years, there remains a need for new treatments, since not all patients respond adequately to current therapies, have co-morbidities and some patients experience toxicities and/or intolerance that limit the use of approved therapies.

In November 2012, tofacitinib (Xeljanz®) became the first Janus kinase (JAK) inhibitor to receive Food and Drug Administration (FDA) approval for the treatment of adult patients with RA. Tofacitinib is a small molecule, has strong binding affinity for JAK1 and JAK3, and weaker affinity for JAK2. The extensive pre-clinical and clinical development programs demonstrated its mechanisms of action via anti-inflammatory and immunosuppressive effects. The drug proved to be efficacious in treating the signs and symptoms of RA. However, the observed side-effects and risk profile of tofacitinib are similar to those of several existing anti-rheumatic agents with cytopenias, elevated levels of liver function enzymes, increased total cholesterol levels, with increase in LDL typically exceeding those for HDL, and increased risk for infections including serious and opportunistic infections. At higher doses, tofacitinib treatment was associated with anemia, which is thought to be linked to inhibition of JAK2.

While the panJAK inhibitor tofacitinib has shown an early onset of action and long-term efficacy in RA as mono therapy and in combination with background conventional synthetic disease modifying anti-rheumatic drugs (csDMARDs) therapy, dose levels were limited by side effects potentially mediated by its effect on JAK 2 and JAK 3. This highlights the need for more selective and targeted therapies with improved immunomodulatory and hematologic effects.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

JAK1 is thought to be an integral part of RA pathogenesis due its role in transmitting inflammatory cytokine signaling. Hence, targeted inhibition of JAK1 has great potential for the treatment of RA with an improved safety and side effect profile.

1.2. Filgotinib

1.2.1. General Information

Janus kinases are intracellular cytoplasmic tyrosine kinases (TY Ks) that transduce cytokine signaling from membrane receptors through signal transducer and activator of transcription (STAT) to the nucleus of cells. JAK inhibitors block the signaling of various cytokines, growth factors, and hormones, including the pro-inflammatory cytokine interleukin (IL)-6.

Four different types of JAKs are known, JAK1, JAK2, JAK3, and TYK2 which co-interact with different sets of membrane receptors. Inhibition of JAKs is a promising therapeutic option for a range of inflammatory conditions including RA and Crohn’s Disease (CD).

Filgotinib (GS-6034, formerly GLPG0634) is a potent and selective inhibitor of JAK1. The compound has shown good preliminary efficacy in RA and CD patients in Phase 2. No typical JAK2 side effects such as anemia were observed in clinical studies of up to 24 weeks duration.

In humans, filgotinib is metabolized to form one major active metabolite GS-829845 (formerly G254445). Though the potency of this metabolite is lower than the parent molecule, the overall exposure and peak plasma concentration in humans is higher than seen in all tested animal species. As a consequence, dedicated pharmacology and toxicology studies have been performed with GS-829845. Results from pharmacodynamics (PD) testing in healthy volunteers suggest that the clinical activity of filgotinib could result from the combination of the parent molecule and the metabolite.

For further information on filgotinib, refer to the current investigator’s brochure (IB).

1.2.2. Nonclinical Pharmacology, Absorption, Distribution, Metabolism, and Excretion (ADME) and Toxicology

Filgotinib and its metabolite, GS-829845 have been extensively characterized in nonclinical studies. This program includes cellular assays demonstrating potency and selectivity of the compound against JAK1; efficacy studies in rats and mice; repeat dose toxicity studies (up to 26 weeks in the rat and 39 weeks in the dog), in vitro and in vivo safety pharmacology and genetic toxicology studies, and reproductive toxicology studies in rats and rabbits. Additional toxicology studies conducted include phototoxicity studies and dose-range finding studies in support of a definitive rat juvenile toxicity study and a 6 month carcinogenicity study in transgenic (TgαH2) mice. A 2 year rat oral carcinogenicity study is ongoing.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

1.2.2.1. Primary and Secondary Pharmacology

Filgotinib is an adenosine triphosphate (ATP)-competitive inhibitor of JAK1. It is highly selective for inhibition of JAK1 over 451 other kinases evaluated in vitro. In cellular assays, it inhibits JAK/STAT-driven processes with half maximal inhibitory concentration (IC₅₀) values from 179 nM onwards. In human whole blood, JAK1 is inhibited by filgotinib with an IC₅₀ of 629 nM and exhibits approximately 30-fold selectivity over JAK2. Filgotinib demonstrated significant efficacy in the rat collagen-induced arthritis (CIA) model as well as in the mouse dextran sulphate sodium (DSS)-induced colitis model.

Metabolite GS-829845 exhibits a similar JAK1 selectivity profile but is approximately 10 to 20-fold less potent than the parent filgotinib in in vitro assays. GS-829845 was as effective as filgotinib in the rat CIA model, but at doses that required a 10-fold higher exposure.

1.2.2.2. Safety Pharmacology

Filgotinib and GS-829845 had no effects on the respiratory system and central nervous system (CNS) up to respectively 40- and 5-fold the exposure in RA subjects given filgotinib 200 mg q.d.

Filgotinib and GS-829845 had no relevant effects on cardiovascular parameters (human ether-a-go-go related gene [hERG] and dog telemetry studies), apart from a slight non-adverse increase in heart rate and arterial pressure with GS-829845 at exposures 7-fold that of the Cₘₚ in subjects with RA treated with 200 mg q.d. filgotinib. There were no relevant effects on electrocardiogram (ECG) and QT.

1.2.2.3. Nonclinical ADME

Filgotinib demonstrates good oral bioavailability in mice, rats, dogs, and minipigs but less in monkeys. Plasma protein binding is low (< 70%) in all species, including humans.

The pharmacokinetics (PK) of filgotonib is generally dose proportional without gender differences. No accumulation occurs with repeated dosing. The mean terminal half-life after oral administration is 4 and 5 hours (h) in rats and dogs, respectively.

In the rat, filgotinib showed a rapid and even distribution throughout the body. High concentrations were observed only in the gastrointestinal (GI) tract and urinary bladder. Filgotinib does not penetrate into central nervous system (CNS) tissues. The distribution of filgotinib indicates some affinity for melanin-containing tissues.

Excretion is nearly complete within 24 h (rat) and 48 h (dog) post-dosing. In the rat, fecal and urinary excretion accounted for 40% and 53% of the administered dose, respectively, with a bile secretion of about 15%. In the dog, fecal excretion was the primary route of excretion, accounting for 59% of the administered dose, with urinary excretion accounting for 25%.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

In vitro metabolism studies in all species revealed one major metabolite (GS-829845). The formation of GS-829845 is mediated by carboxylesterases (CES) and is not dependent on cytochrome P450 (CYP).

In vitro experiments have shown that drug-drug interactions with filgotinib and GS-829845 are unlikely. There is no inhibition or induction of CYPs or uridine 5’-diphosphate glucuronosyltransferases (UGTs), and no relevant inhibition of key drug transporters, including the organic anion transporters (OATs) involved in the renal elimination of MTX, by filgotinib or GS-829845.

1.2.2.4. Nonclinical Toxicology

In repeat oral dose toxicity studies in both rats and dogs, the primary target tissues identified for filgotinib were the lymphoid tissues which is expected based on the pharmacology of JAK inhibition. Additional filgotinib-related findings were observed in the male reproductive organs of both species, and in the incisor teeth of rats only. Effects on the lymphoid system were fully reversible. Testicular toxicity demonstrated partial reversibility, however sperm counts remained low. A dose of 200 mg/day of filgotinib results in an estimated mean clinical AUC of 2.80 μg/mL, which represents an exposure margin of 2.3, 1.8, and 3.4-fold when considering the mean AUC in male dogs at the no-observed-effect-levels (NOELs) in the 26 week and 39 week chronic toxicity studies, and the 39 week targeted exposure toxicity study, respectively.

GS-829845-related findings in general repeat dose toxicity studies were similar to those of the parent filgotinib, however no testicular toxicity was noted following administration of GS-829845.

Filgotinib and GS-829845 were non-genotoxic when evaluated in the bacterial mutagenicity assay, the in vitro mouse lymphoma mutagenicity assay, and the rat bone marrow micronucleus assay.

In embryofetal development studies, filgotinib and GS-829845 caused embryolethality and teratogenicity in rats and rabbits at exposures similar to the human exposure at 200 mg q.d. of filgotinib in subjects with RA. A administration of filgotinib did not affect female fertility but impaired fertility was observed in male rats at exposures approximately 15-fold the human exposure at 200 mg of filgotinib in subjects with RA. GS-829845 did not have any effects on fertility parameters in either male or female rats.

In an in vitro phototoxicity study in 3T3 cells, the metabolite GS-829845 was positive for phototoxic potential and results with filgotinib were equivocal. A follow-up in vivo rat phototoxicity assay revealed a lack of phototoxic potential for both compounds.

1.2.3. Clinical Trials of Filgotinib

Comprehensive data from the Phase 1 and 2 programs are available to support development into Phase 3. As of January 2016, filgotinib has been administered to more than 150 healthy subjects, more than 1000 RA subjects, and more than 150 subjects with CD. A detailed description of all clinical studies can be found in the IB.
Phase 2b G LPG 0634-CL-203, filgotinib with MTX in RA

In GLPG0634-CL-203, subjects with active RA on stable dose of MTX were randomized to receive either placebo or one of three total daily doses of filgotinib (50 mg, 100 mg, or 200 mg) on a once or twice daily schedule for 24 weeks. The primary objective of the study was to evaluate the efficacy of different doses and dose regimens of filgotinib compared to placebo at Week 12.

The percentage of American College of Rheumatology (ACR) 20 responders was statistically significantly higher in the 100 mg and 200 mg once daily, and 100 mg twice daily dose groups at Week 12 and in the 100 mg and 200 mg once daily, and 50 mg and 100 mg twice daily dose groups at Week 24. The percentage of ACR50 responders was statistically significantly higher compared with placebo across all filgotinib dose groups and regimens at both Weeks 12 and 24 (Table 1-1). The percentage of ACR70 responders was statistically significantly higher in the filgotinib 200 mg once daily and 100 mg twice daily dose groups compared with placebo at Week 12 and across all filgotinib dose groups and regimens at Week 24. A dose-response was observed for all three parameters. In addition, the ACR20 response appeared to plateau at Week 8 in the majority of filgotinib treatment groups and was maintained up to Week 24. At Week 24, the ACR50 response was maintained and the ACR70 response continued to increase compared with Week 12.

Starting at week 2 response was observed for ACR20 and ACR50. No statistically significant difference was found between the once and twice daily regimens.

Table 1-1. Summary and analysis of ACR20/50/70 response at Weeks 12 and 24 (NRI [ITT Population]), GLPG0634 - CL-203

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Time Point</th>
<th>Placebo N=86</th>
<th>Filgotinib once daily Dose Groups</th>
<th>Filgotinib twice daily Dose Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N=82</td>
<td>50 mg N=82</td>
<td>100 mg N=85</td>
</tr>
<tr>
<td>ACR20</td>
<td>W12</td>
<td>38 (44.2)</td>
<td>46 (56.1)*</td>
<td>54 (63.5)*</td>
</tr>
<tr>
<td></td>
<td>W24</td>
<td>36 (41.9)</td>
<td>45 (54.9)*</td>
<td>52 (61.3)*</td>
</tr>
<tr>
<td>ACR50</td>
<td>W12</td>
<td>13 (15.1)</td>
<td>27 (32.9)*</td>
<td>32 (37.6)**</td>
</tr>
<tr>
<td></td>
<td>W24</td>
<td>14 (16.3)</td>
<td>29 (35.4)**</td>
<td>40 (47.1)**</td>
</tr>
<tr>
<td>ACR70</td>
<td>W12</td>
<td>7 (8.1)</td>
<td>13 (15.9)</td>
<td>18 (21.2)</td>
</tr>
<tr>
<td></td>
<td>W24</td>
<td>8 (9.3)</td>
<td>18 (22.0)*</td>
<td>28 (32.9)**</td>
</tr>
</tbody>
</table>

Note 1: p-values were based on a pairwise comparisons of each group vs. the placebo group using a logistic regression model with factors treatment group, geographical region, and prior use of biologics. Hommel-corrected p-value. Note 2: The denominator for the percentage calculations = the total number of subjects per group with a response (yes or no) at that time point.
Note 3: Subjects who switched treatment at Week 12 were handled as if they discontinued at Week 12.
* p < 0.05; ** p < 0.01; *** p < 0.001
ACR = American College of Rheumatology; ITT = Intent-to-treat; NRI = non-responder imputation; W = week
Source: GLPG0634-CL-203

Filgotinib  
Protocol GS-US-417-0301  
Gilead Sciences, Inc.

At Weeks 12 and 24 the mean decrease in Disease Activity Score for 28 joint count using c-reactive protein (DAS28[CRP]) was statistically significantly greater across all filgotinib dose groups and regimens compared with placebo. A dose-response was observed. No statistically significant difference was apparent between the once and twice daily regimens. At Weeks 12 and 24, the mean decrease in Simplified Diagnostic Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) score was statistically significantly greater across all filgotinib dose groups and regimens compared with placebo (with the exception of filgotinib 50 mg once daily dose group at Week 12). In addition, the mean SDAI and CDAI scores were maintained after Week 12 in the 50 mg daily dose groups and continued to improve up to Week 24 in the 100 mg and 200 mg daily dose groups.

No unexpected safety findings were noted. Overall, no differences were observed in the incidence of treatment emergent adverse events (TEAEs) reported for subjects in any of the dosing groups, including placebo, for the duration of the study. TEAEs were reported for 51.2% of ˚All Placebo Exposed_ subjects (ie, all subjects combined who received placebo during either the entire 24 weeks or only during the first 12 weeks) and 51.5% of ˚All filgotinib Exposed_ subjects (ie, all subjects combined who received filgotinib during either the entire 24 weeks or only during the last 12 weeks, irrespective of dose).

A total of 15 subjects had h1 serious TEAE; 4 subjects in the placebo group (4.7%) and 11 subjects (2.0%) in one of the filgotinib groups. One of these subjects with h1 serious TEAE, who received filgotinib 100 mg twice daily with concurrent MTX, died during the second 12 weeks of the study period due to pneumonia and septic shock. Out of the 15 subjects with a serious TEAE, 11 subjects had a serious TEAE due to which the study medication was stopped and the subject discontinued the study. A total of 23 subjects had h1 AE leading to permanent discontinuation of the study medication and the study; 2 subjects (2.3%) in the placebo group and 21 subjects (3.9%) in one of the filgotinib groups (including a subject in the filgotinib 100 mg q.d. group who had a pre-dosing AE which was ongoing throughout the study, for which the study medication was permanently discontinued). Most of the serious TEAEs and the AEs leading to discontinuation (by preferred term) were experienced by a single subject.

For the duration of the study, the most common (h10%) TEAEs reported by SOC in subjects from both the placebo and filgotinib dosing groups, were Infections and Infestations and Gastrointestinal disorders. There were no differences between subjects who received placebo or filgotinib in the severity of TEAEs (most TEAEs were mild or moderate; severe TEAEs were observed in 1.2% of ˚All Placebo Exposed_ subjects and in 2.2% of ˚All filgotinib Exposed_ subjects). Treatment-related TEAEs were generally reported more often for subjects in the filgotinib dosing groups than in the placebo group (9.3% with placebo and 20.3% with filgotinib); however, within the different filgotinib dosing groups, no clear dose relationships were observed.

Six serious infections were reported (1 in placebo arm, 5 in filgotinib arm). All 6 serious and one additional non-serious infection in the filgotinib group led to dosing discontinuation. Up to Week 24, herpes zoster infections were observed in 5 subjects (1 placebo treated patient and 4 filgotinib). No cases of tuberculosis, opportunistic infections, lymphoma, or cancer were reported throughout the 24-week dosing period.

CONFIDENTIAL  
Page 24  
05 July 2016
Laboratory data were consistent with prior phase 2 studies and no new safety findings were observed from laboratory data. A summary of laboratory findings of interest, including hemoglobin, neutrophil, lymphocyte, creatinine, lipid, and hormone data are summarized below.

Up to Week 12, small increases were observed in mean hemoglobin concentrations in the filgotinib 200 mg daily dose groups (increase of 4.4 g/L from baseline in the filgotinib 100 mg bid group). Thereafter, hemoglobin mean concentrations appeared to plateau and remain stable until Week 24 (increase of 4.9 g/L from baseline in the filgotinib 100 mg bid group).

Up to Week 4, dose-dependent decreases were observed in mean absolute neutrophil counts in the filgotinib treatment groups. Mean absolute neutrophil counts appeared to plateau and remained stable until Week 24. No decreases in mean absolute lymphocyte counts were observed, including lymphocyte subsets.

Up to Week 4, dose-dependent decreases were observed in mean absolute platelet counts in the filgotinib treatment groups. Mean absolute platelet counts appeared to plateau and remained stable. Dose-dependent increases in the filgotinib groups were observed in mean creatinine concentrations during the first 4 weeks of the study for most filgotinib treatment groups (up to Week 8 for the filgotinib 100 mg bid group) that subsequently plateaued and remained stable up to Week 24.

Up to Week 4, dose-dependent increases were observed in mean concentrations of total cholesterol, LDL cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides in all filgotinib treatment groups. All these lipid parameters further increased up to Week 8 in the filgotinib 200 mg daily dose groups. Thereafter, these increases appeared to plateau and maintained at stable mean concentrations up to Week 24. At Week 24, mean increases were observed of 0.7 mmol/L in total cholesterol, 0.3 mmol/L in LDL cholesterol, 0.3 mmol/L in HDL cholesterol, and 0.1 mmol/L in triglycerides in the filgotinib 100 mg bid group.

In male subjects, small non dose-dependent increases were observed in total and free testosterone during the study (at Week 24, mean increases were 3.4 nmol/L for total and 51.7 pmol/L for free testosterone in the filgotinib 100 mg bid group). For FSH, inhibin B, LH, and prolactin, small changes (both increases and decreases) were observed during the study, without any trends of larger changes in male subjects of one or more of the different treatment groups.

**GLPG 0634-CL-204, Filgotinib administered as monotherapy in RA Subjects**

The primary objective of study GLPG0634-CL-204 was to evaluate the efficacy of three doses of filgotinib q.d. compared to placebo at Week 12.

As shown in Table 1-2, the percentage of ACR20 and ACR50 responders at week 12 was statistically significantly higher across all filgotinib dose groups compared with placebo. The percentage of ACR70 responders in the filgotinib 100 mg and 200 mg once daily dose groups was statistically significantly higher compared with placebo. At Week 24, the ACR50 response was maintained and the ACR70 response showed continued improvement. An early onset of response was observed for ACR20 (from Week 1 in the filgotinib 200 mg once daily...
dose group and Week 4 across all other dose groups), ACR50 (from Week 2 in the filgotinib 200 mg once daily dose group and Week 4 across all other filgotinib dose groups), and ACR70 (Week 4 in the filgotinib 200 mg once daily dose group). The time to ACR20/50/70 response was shorter in all filgotinib dose groups compared with placebo.

Table 1-2. Summary and analysis of ACR20/50/70 response at Weeks 12 and 24 (NRI [ITT Population]); GLPG0634-0204

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Time Point</th>
<th>Placebo</th>
<th>Filgotinib once daily Dose Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=72</td>
<td>50 mg N=72</td>
<td>100 mg N=70</td>
</tr>
<tr>
<td>ACR20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W12</td>
<td>21 (29.2)</td>
<td>48 (66.7)***</td>
<td>46 (65.7)***</td>
</tr>
<tr>
<td>W24</td>
<td>Not applicable</td>
<td>41 (56.9)</td>
<td>55 (78.6)</td>
</tr>
<tr>
<td>ACR50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W12</td>
<td>8 (11.1)</td>
<td>25 (34.7)***</td>
<td>26 (37.1)***</td>
</tr>
<tr>
<td>W24</td>
<td>Not applicable</td>
<td>24 (33.3)</td>
<td>27 (38.6)</td>
</tr>
<tr>
<td>ACR70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W12</td>
<td>2 (2.8)</td>
<td>6 (8.3)</td>
<td>13 (18.6)**</td>
</tr>
<tr>
<td>W24</td>
<td>Not applicable</td>
<td>14 (19.4)</td>
<td>18 (25.7)</td>
</tr>
</tbody>
</table>

Note 1: p-values were based on a pairwise comparisons of each group vs. the placebo group using a logistic regression model with factors: treatment group, geographical region, and prior use of biologics. Hommel-corrected p-value.

Note 2: The denominator for the percentage calculations = the total number of subjects per group with a response (yes or no) at that time point

Note 3: Subjects who switched treatment at Week 12 were handled as if they discontinued at Week 12.

* p<0.05; ** p<0.01; *** p<0.001

ITT=Intent-to-treat; NRI=non-responder imputation; W=week

Source: GLPG0634-CL-204

At Week 12, the mean decrease in DAS28(CRP) was statistically significantly greater across all filgotinib dose groups compared with placebo. At Week 24, the mean decrease in DAS28(CRP) was maintained in the 50 mg once daily dose group and showed a small improvement in the highest dose groups. In addition, at Week 12, the percentage of subjects with DAS28(CRP) remission was higher across all filgotinib dose groups compared with placebo. Differences vs. placebo were not statistically significant for any of the filgotinib dose groups. The number of subjects with DAS28(CRP) < 2.6 and < 3.2 were higher across all filgotinib dose groups compared with placebo at Week 12; differences vs. placebo were statistically significant for the filgotinib 200 mg once daily dose group.

Safety data revealed no differences in the incidence of TEAEs reported for subjects in any of the treatment groups, including placebo, during both the first 12 weeks of treatment and the full 24 weeks of treatment. TEAEs were reported for 38.9% of "All Placebo Exposed" subjects (ie, all subjects combined who received placebo during the first 12 weeks) and 41.3% of "All filgotinib Exposed" subjects (ie, all subjects combined who received filgotinib during either the entire 24 weeks or only during the last 12 weeks, irrespective of dose).
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Final Amendment 1

No deaths were reported and a total of 9 subjects had a serious TEAE; 1 subject (1.4%) during placebo dosing and 8 subjects (2.9%) during filgotinib dosing. No serious TEAE (by preferred term) was experienced by more than 1 subject, and all subjects recovered from their serious TEAEs. Out of the 9 subjects with a serious TEAE, 3 subjects had a serious TEAE for which the study medication was stopped and the subject discontinued the study. There were no differences in incidences of AEs leading to discontinuation among all the different dosing groups, including placebo. A total of 11 subjects had ≥1 TEAE leading to discontinuation of the study medication; 4 subjects (5.6%) during placebo dosing and 7 subjects (2.5%) during filgotinib dosing.

During the whole study, the most common TEAEs reported by System Organ Class in subjects from both the placebo and filgotinib treatment groups, were: Infections and Infestations and Gastrointestinal disorders. There were no differences between subjects who received placebo or filgotinib in the severity of TEAEs (most TEAEs were mild or moderate; severe TEAEs were observed in 1.4% of All Placebo Exposed subjects and in 1.1% of All filgotinib Exposed subjects). Treatment-related TEAEs were generally reported more often for subjects in the filgotinib treatment groups than in the placebo group (9.7% with placebo and 16.7% with filgotinib); however, within the different filgotinib treatment group, no clear dose relationships were observed.

Low numbers of infections were reported as serious (4 subjects with filgotinib) or led to discontinuation of the study medication (2 serious infections; ie, cellulitis and pneumonia) were observed during the study. Up to Week 24, 1 subject (filgotinib 50 mg q.d. group) had a herpes zoster infection. No cases of tuberculosis, opportunistic infections, lymphoma, or cancer were reported throughout the 24-week treatment period.

Laboratory data were consistent with prior studies and no new safety findings were observed.

Please refer to the IB for additional data regarding efficacy and safety.

1.3. Information about Humira® (adalimumab)

For information on adalimumab, refer to the current package insert.

1.4. Rationale for this Study

Over the last decade, changes in RA treatment strategies, accompanied by advances in drug development and the addition of targeted biological therapies, have greatly improved the outcomes for subjects with RA. Despite these developments, therapeutic challenges remain. The current conventional and biologic disease-modifying anti-rheumatic drugs (bDMARDs) may be ineffective or produce only partial responses in some subjects and may be associated with significant safety and tolerability concerns. There is a medical need for simple, orally administered therapies with novel and targeted mechanisms of action that can effectively improve the disease course while being safe and well-tolerated.
Filgotinib is an orally administered, small molecule inhibitor of JAK1, an intracellular tyrosine kinase dysregulated in subjects with inflammatory disorders including RA. Filgotinib has demonstrated clinical activity and a favorable safety and tolerability profile in Phase 2 studies in subjects with moderately to severely active RA.

1.4.1. Rationale for Study Design

GS-US-417-0301 is a Phase 3 randomized, double-blind, placebo-controlled and active-comparator study designed to evaluate the efficacy and safety of filgotinib in subjects with active RA who have an inadequate response to MTX therapy (MTX-IR). Filgotinib will be administered orally for up to 52 weeks. The objectives of the study are to evaluate the efficacy of filgotinib on: the signs and symptoms of RA, on physical function, radiographic progression, work productivity, fatigue, and quality of life. In addition, the safety, tolerability, and pharmacokinetics (PK) of filgotinib will be assessed.

All subjects in this study are required to maintain background therapy with MTX. Methotrexate is the most commonly prescribed first line csDMARD in the treatment of moderately to severely active RA and has been shown to reduce the signs and symptoms of disease as well as to inhibit radiographic progression {Lopez-Olivo et al 2014}. The placebo control used in this study will allow assessment of the filgotinib safety and tolerability profile and establish the efficacy contribution that filgotinib may provide in combination with MTX. Nonresponders at Week 14, as defined in protocol Section 3.2, will discontinue investigational therapy but will continue with study visits and assessments per protocol. All subjects who discontinue from investigational therapy, are to receive standard of care treatment for their RA as determined by the investigator.

The duration of placebo will be limited. All subjects in the placebo arm will be reallocated 1:1 to 100 mg or 200 mg of filgotinib at Week 24. Reallocation of treatment assignment at Week 24 will be performed in a blinded manner. Given the slowly progressive nature of RA, 14 weeks of placebo in Non-responders on accepted background therapy is considered both clinically and ethically appropriate and allows an acceptable time period in which to assess initial response to therapy. The use of placebo is critical in evaluating subject reported and subjective outcomes, which can be highly variable and influenced by placebo therapy. The placebo effect is well-established in RA studies and inclusion of the placebo arm is critical to evaluation of the overall risk:benefit profile of added filgotinib therapy. Finally, the provision of best supportive care in the context of the trial, as well as the limited duration and eventual opportunity for subjects to transition to filgotinib, make the use of placebo ethically acceptable for the specified duration.

This study includes an active comparator, adalimumab, a monoclonal antibody directed against tumor necrosis factor (TNF-). The inclusion of the active comparator will allow comparison of the efficacy and safety of filgotinib to an approved and frequently prescribed product indicated for use in patients who have an inadequate response to MTX therapy.

The study eligibility criteria are consistent with those of recent clinical trials evaluating novel investigative treatments for RA. Subjects should meet the 2010 ACR/European League Against Rheumatism (EULAR) criteria for RA. Subjects are required to have active disease.
defined as ≥6 swollen and ≥6 tender joints at both Screening and Day 1. In addition, subjects are required to have one of the following: an elevated CRP, or the presence of either 3 or more documented joint erosions on radiographs if seronegative, or 1 documented joint erosions on radiographs if seropositive for anti-CCP antibody or RF. The study duration of up to 52 weeks allows sufficient time for demonstrating the safety and tolerability of therapy, reduction in disease activity, and confirmation that clinical benefit is sustained over time.

1.4.2. Rationale for the Outcome Measures

Safety and tolerability will be assessed by the evaluation of adverse events (AEs), selected clinical laboratory parameters, vital signs, physical examinations, and ECGs, all of which are standard safety evaluations in clinical research studies.

The ACR20, ACR50, and ACR70 responses and the DAS28 (CRP) are considered reliable measures of response to treatment and disease activity, respectively in subjects with RA. Comparison between the treatment groups of the proportions of subjects achieving an ACR20 response at Week 12 and the proportions of subjects achieving low disease activity in DAS28 (CRP) allows for interpretation of a clinically meaningful response and these measures have been shown to achieve high discriminatory capacity.

Evaluation of the continuous outcome measures of DAS28 (CRP) and American College of Rheumatology N (ACR-N) as secondary endpoints enable the demonstration of improvement and magnitude of benefit. The EULAR response criteria classify subjects as non-, moderate-, or good responders depending on the extent of change and the level of disease activity reached. These response criteria are useful when describing clinically meaningful therapeutic results. The CDAI and the SDAI have been widely used in clinical studies to demonstrate the impact of a study drug on controlling disease activity in patients with RA.

Assessing quality of life (measured by the Functional Assessment of Chronic Illness Therapy-Fatigue [FACIT-Fatigue], 36-item short form health survey [SF-36], and EuroQol 5-Dimension [EQ-5D]) at Day 1 and during the course of study treatment provides insight into the effects on modifying RA disease course and impact on daily life. Treatment satisfaction will be assessed by using the Treatment Satisfaction Questionnaire for Medication (TSQM) at Day 1 and during the course of the study. Assessing the change in economic impact (Work Productivity and Activity Impairment [WPAI] and Healthcare Resource Utilization) of RA over the course of the study provides insight into the subject’s ability to work and other daily activities as well as the impact on the burden of healthcare resources.

Radiographs have been widely used in assessing structural joint damage associated with RA and are therefore essential in evaluating the efficacy of a study drug. Reduction in radiographic evidence of structural damage progression is an important predictor of long-term benefits in delaying or preventing the progression to disability related to RA. While X-ray remains the widely accepted method for assessing progression of structural joint damage, evaluation by MRI may provide specific advantages and an exploratory assessment of structural joint damage by MRI will also be made during the study. Radiologic data using validated scoring methods will be used to demonstrate efficacy in this domain.
1.4.3. Rationale for the Choice of Dose and Dosing Interval

Enrolled subjects will be randomized to receive filgotinib (100 mg or 200 mg), active comparator (adalimumab), or placebo. The 100 mg and 200 mg q.d. dose regimens of filgotinib are based on efficacy and safety data from the Phase 1 and 2 studies, and are supported by non-clinical safety studies.

In seven Phase 1 studies conducted in healthy subjects (GPLG0634-CL-101, -102, -103, -104, -105, -107, and -110), filgotinib administered at doses up to 450 mg q.d. for up to 10 days were safe and well tolerated.

In the two Phase 2a studies in subjects with RA (Study GPLG0634-CL-201 and -202), dosing with filgotinib was well tolerated and achieved a high level of efficacy at a 200 mg daily dose (ACR20 response of 75-92% at Week 4). Administration of a higher filgotinib dose (300 mg) did not demonstrate greater efficacy, therefore, the highest dose to be tested in this study will be 200 mg q.d.

In two Phase 2b studies, filgotinib at total daily doses of 50 mg, 100 mg, or 200 mg, administered in addition to a background therapy with MTX (GPLG0634-CL-203) or as monotherapy (GPLG0634-CL-204) was shown to be safe and efficacious in subjects with moderately to severely active RA who had an inadequate response to MTX alone (Section 1.2.3).

Exposure-response analysis based on data from all Phase 2 studies indicated a dose-dependent increase in efficacy (ACR20/50/70, DAS28[CRP]), with a plateau at the 200 mg total daily dose on the dose-response curve. Additionally, in Study GPLG0634-CL-203, no statistically significant difference in efficacy was observed at 200 mg daily dose, administered as 200 mg q.d. or 100 mg b.i.d. These results are consistent with the relationship observed between filgotinib exposures and pSTAT1 activation (ex-vivo) following single and multiple filgotinib doses, where maximal inhibition of pSTAT1 (~78%) was achieved at or above 200 mg total daily dose and intermediate inhibition (~47%) at 100 mg (Namer et al 2015).

Safety data collected across Phase 2 clinical studies showed no dose-dependent trends in the incidence of AEs or SAEs, including infections, or laboratory abnormalities with the exception of a numerical increase in select gastrointestinal AEs (eg, nausea, vomiting, abdominal pain, and upper abdominal pain). This numerical increase was observed in the 200 mg compared to the 100 mg dose. However, the overall frequency was low and clinical relevance is unknown. Filgotinib, administered at a dose of 100 mg or 200 mg daily was found to be safe and well tolerated. The safety profile was consistent with that observed for an immunomodulatory compound administered to subjects with RA.

Overall, the 100 mg and 200 mg once-daily dose regimens have been proposed based on the safety and efficacy data from the Phase 2 studies in RA and the observed plateau in the pSTAT1 response which indicates doses above 200 mg are unlikely to add additional benefit. Inclusion of multiple doses in the proposed Phase 3 trials will enable establishment of an appropriate nominal dose for the treatment of RA and determine the regimen with the most favorable risk:benefit profile in these populations.
1.5. Risk/Benefit Assessment for the Study

As of January 2016, Filgotinib has been administered to more than 150 healthy subjects as single doses ranging from 1 to 200 mg, and as multiple doses at 25 to 100 mg twice daily, and 200, 300, and 450 mg daily for 10 days. In addition, filgotinib has been administered to more than 1000 RA subjects and more than 150 subjects with Crohn’s Disease at daily doses ranging from 50-200 mg. In general, filgotinib has been safe and well tolerated in all populations studied.

Nonclinical studies in rats and dogs identified the testes and lymphoid tissue as target organs for filgotinib in long term repeat-dose toxicity studies. In both species, histopathological changes in the testes included germ cell depletion and degeneration, with reduced sperm content and increased cell debris in the epididymis and reduction in fertility in rats. The dog was determined to be the most sensitive species. A dose of 200 mg/day of filgotinib results in an estimated mean clinical AUC of 2.8 \( \mu g/mL \), which represents an exposure margin of 2.3, 1.8, and 3.4-fold when considering the mean AUC in male dogs at the no-observed-effect-levels (NOELs) in the 26 week and 39 week chronic toxicity studies, and the 39 week targeted exposure toxicity study, respectively. Decreased lymphocytes observed in nonclinical studies have not been shown in clinical studies.

Filgotinib has shown an increased risk of embryofetal malformations at exposures similar to human doses; the use of highly effective contraception in the subject population will be implemented in the study to mitigate this risk.

No clinically relevant impact on cardiovascular parameters (including vital signs and ECG), respiratory or neurologic function has been observed in Phase 1 and 2 trials of filgotinib. Across phase 2 trials in RA, filgotinib was well tolerated. In the RA studies (including the open label extension Darwin 3), infections were reported more commonly in the filgotinib groups, including serious infections leading to hospitalization, and even death. The most common system organ classes (SOC) with AEs were infections and infestations, and gastrointestinal disorders. Dose dependent decreases in the phase 2b studies were observed in mean neutrophil counts and platelet counts (but mean changes in both remained within normal laboratory reference ranges), and there were no decreases in lymphocytes or lymphocyte subsets. Hemoglobin levels slightly improved (increased) with filgotinib treatment, confirming that no anemia was induced. Mild and clinically insignificant serum creatinine increases were noted in both Phase 2b studies, with stabilization by Week 24. Neutrophil decreases (in the RA population) and a potential increased risk of infection may be considered risks consistent with the mechanism of JAK inhibition.

Overall clinical findings and laboratory changes are consistent with selective JAK 1 inhibition and based on Phase 2 data the expected benefit of using filgotinib as proposed in this study is considered to outweigh any associated risks.

Adalimumab is a TNF inhibitor that is indicated and commonly used in the target population of this trial with well established safety and efficacy profile (Humira US PI, Humira EU SmPC).

The dose and duration of adalimumab treatment proposed in this study are in line with posology approved globally (Humira US PI, Humira EU SmPC).
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Risks associated with the use of adalimumab are well characterized, and are described in approved local product information. Adalimumab is associated with an increased risk for developing serious infections that may lead to hospitalization or death. Opportunistic infections due to bacterial, mycobacterial, invasive fungal, viral, parasitic, or other opportunistic pathogens have been reported with TNF inhibitors. Reactivation of latent TB has been reported. All subjects will be tested for TB before study start as outlined in the exclusion criteria and will be carefully monitored for signs and symptoms of infection during and after the study. Study drug will be discontinued for serious infections as outlined in section 3.5. An increase of non-melanoma skin cancers, hematologic cancers and a potential increase of other cancers has been reported with use of TNF inhibitors. Subjects will undergo regular assessments, including physical examination and haematological assessments.

TNF inhibitors are commonly used as first line therapy in MTX-IR RA patients. Their efficacy in this patient population has been well established. Overall, the benefit of using adalimumab as proposed in this study is considered to outweigh any associated risks.

An independent data monitoring committee (DMC) appointed to monitor the study (with an interim safety analysis after the first 100 subjects complete 12 weeks of treatment) will provide an additional level of risk mitigation.

The overall risk:benefit balance of this study is considered favorable. For additional information about the risks of filgotinib, reference is made to the investigator brochure.

1.6. Compliance

This study will be conducted in compliance with this protocol, Good Clinical Practice (GCP), and all applicable regulatory requirements.
2. **OBJECTIVES**

The primary objective of this study is as follows:

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving ACR20 at Week 12

The secondary objectives of this study are as follows:

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving Disease Activity Score for 28 joint count using CRP (DAS28 [CRP]) ≥3.2 at Week 12

- To evaluate the effect of filgotinib versus placebo on physical function as measured by change from Baseline in the Health Assessment Questionnaire Disability Index (HAQ-DI) score at Week 12

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP) <2.6 at Week 24

- To evaluate the effects of filgotinib versus placebo on preservation of joint structure as measured by change from Baseline in the van der Heijde modified total Sharp score (mTSS) at Week 24

- To evaluate the effects of filgotinib versus adalimumab for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP) ≥3.2 at Week 12

- To evaluate the safety and tolerability of filgotinib

- To evaluate the effects of filgotinib on work productivity, fatigue, and general quality of life as measured by SF-36, FACIT-Fatigue, EQ-5D and WPAI-RA

The exploratory objectives of this study include:

- To characterize the pharmacokinetics (PK) of filgotinib and its metabolite (GS-829845)

- To characterize the association of host genetics and other markers with disease severity, disease progression and treatment response to filgotinib in subjects with rheumatoid arthritis

- To evaluate the effects of filgotinib on healthcare resource utilization and other patient reported outcomes
3. STUDY DESIGN

3.1. Endpoints

The primary endpoint is:

- The proportion of subjects who achieve an ACR20 response at Week 12
- The key secondary endpoints are:
  - The proportion of subjects who achieve DAS28 (CRP)Hβ.2 at Week 12
  - Change from Baseline in the HAQ-DI score at Week 12
  - The proportion of subjects who achieve DAS28 (CRP)<2.6 at Week 24
  - Change from Baseline in mTSS at Week 24

Other secondary endpoints include:

- The proportion of subjects who achieve ACR50 and ACR70 at Weeks 4, 12, 24 and 52, ACR20 at Weeks 4, 24, and 52, and ACR20/50/70 over time from Day 1 through Week 52
- Change from Baseline in individual components of the ACR response at Weeks 4, 12, 24, and 52 and over time from Day 1 through Week 52
- The proportion of subjects who achieve change in HAQ-DI of h0.22 at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in DAS28 (CRP) at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28 (CRP)Hβ.2 at Weeks 4, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28 (CRP)<2.6 at Weeks 4, 12, and 52, and over time from Day 1 through Week 52
- ACR-N and EULAR response at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in CDAI at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
• Change from Baseline in SDAI at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 24

• Change from Baseline in the mTSS at Week 52

• The proportion of subjects with no radiographic progression from Baseline at Week 24 and 52

• Absolute value and change from Baseline in SF-36, FACIT-Fatigue, and the EQ-5D over time at Weeks 4, 12, 24 and 52, and over time from Day 1 through Week 52

• Absolute value and change from Baseline in WPAI-RA at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52

3.2. Study Design

This is a randomized, double-blind, placebo- and active-controlled, Phase 3 study in adult male and female subjects with active RA who have an inadequate response to MTX. The study is designed to evaluate the efficacy, safety and tolerability of filgotinib as well as its effect on patient-reported outcomes, including work productivity, fatigue, and quality of life. In addition, PK will be assessed.

Adult male and female subjects with active RA will be screened to determine eligibility as per the inclusion and exclusion criteria (see Section 4.2 and 4.3, respectively). The Screening period will be up to 28 days. The screening window may be extended to up to 42 days prior to the Day 1 visit for subjects who require repeat collection of radiographs.

Written informed consent must be obtained before any study-related procedures take place. During the Screening period, following signing of the informed consent form (ICF), radiographs (hands, wrists, and feet) will be sent for central review for confirmation of the number of documented joint erosions. Note that radiographs need to be sent for central review as soon as possible after capture and prior to Day 1 to allow for central reading and confirmation of the joint erosions for eligibility purposes.

Subjects will be randomized in a 3:3:2:3 ratio to filgotinib 200 mg, filgotinib 100 mg, active comparator (adalimumab), or matching placebo controls administered for up to 52 weeks in a double blind fashion.

At Week 14, subjects who have not achieved a 20% improvement from Day 1 in both swollen joint count (SJC) and tender joint count (TJC) will discontinue investigational therapy but will continue with study visits and assessments per protocol. All subjects who discontinue from investigational therapy, are to receive standard of care treatment for their RA as determined by the investigator. At Week 24 all subjects assigned to placebo + MTX will be reassigned 1:1 to either filgotinib 100 mg q.d. or 200 mg q.d. in addition to MTX in a blinded fashion and will continue the study through Week 52.
All subjects will be evaluated for loss of therapeutic response from Week 30 through the Week 52 study visit. Subjects failing to maintain at least a 20% improvement from Day 1 in TJC and SJC (as confirmed at 2 consecutive visits) will discontinue investigational therapy but will continue with study visits and assessments. All subjects who discontinue from investigational therapy are to receive standard of care treatment for their RA as determined by the investigator.

Subjects who have received at least one dose of study drug and choose to prematurely terminate study participation for any reason should complete an ET visit at the time of study discontinuation, as well as a post dosing visit four weeks after the last dose of study drug (post treatment Week 4) regardless of dosing duration.

At completion of the 52-week treatment period, subjects (who did not discontinue assigned study drug or have not met criteria for loss of therapeutic response) will be provided the option to enter a Long Term Extension (LTE) study (GS-US-417-0304).

Subjects or trained caregivers will self-administer the adalimumab (or adalimumab placebo) injections and will not be asked to return to the clinical study center specifically to receive these injections. Subject (or caregiver) training on this self-injection will take place at Day 1 and can be repeated as required.

For those subjects not entering the LTE, a follow-up visit will be completed at Post Treatment Week 4.

To enhance the safety monitoring during the study, a data monitoring committee (DMC) consisting of independent experts will be convened to periodically review the accumulating safety data for the study. In addition, a cardiovascular safety endpoint adjudication committee (CV-SEAC) will be convened to periodically review and adjudicate all possible major cardiovascular events (MACE).

The assessments planned to be performed at each visit are detailed in the study procedures table (Appendix 2). A schedule of the study design is provided below.
Figure 3-1. Study Design

3.3. Study Treatments

Approximately 1650 subjects will be randomized in a 3:3:2:3 ratio to filgotinib 200 mg, filgotinib 100 mg, active comparator (adalimumab), or matching placebo controls administered for up to 52 weeks:

- Filgotinib 200 mg group: filgotinib (200 mg q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
- Filgotinib 100 mg group: filgotinib (100 mg q.d.) + PTM filgotinib 200 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
- Active comparator group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + adalimumab (40 mg s.c. injection q2w) (N=300)
- Placebo control group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)

NOTE: on study visit days, subjects should wait to take their regularly scheduled dose of study drug until instructed by site personnel, in case there are pre-dose blood draws or other pre-dose procedures required on that day
3.4. **Duration of Treatment**

Subjects are planned to participate in the study for approximately 60 weeks (from Screening visit to Follow-up visit or entry into the LTE); the duration of dosing is up to 52 weeks.

3.5. **Criteria for Interruption or Discontinuation of Study Treatment**

3.5.1. **Study drug interruption considerations:**

The Gilead Medical Monitor should be consulted prior to study drug interruption when medically feasible.

Study drug interruption should be considered in the following circumstances; *prior to resumption of study drug, the investigator should discuss the case with the Gilead medical monitor*:

- Intercurrent illness that would, in the judgment of the investigator, affect assessments of clinical status to a significant degree.
- Subject is scheduled for elective or emergency surgery (excluding minor skin procedures under local or no anesthetics); timing of study drug pausing should be determined in consultation with the Gilead medical monitor.
- Any subject who develops a new infection during the study should undergo prompt and complete diagnostic testing appropriate for an immunocompromised individual, and the subject should be closely monitored.

*NOTE: During the time of study drug interruption for any of the above, the subject may continue to have study visits and to take part in procedures and assessments, if deemed medically appropriate by the investigator.*

3.5.2. **Study drug discontinuation considerations:**

The Gilead Medical Monitor should be consulted prior to study drug discontinuation when medically feasible.

Study medication should be permanently discontinued in the following instances:

- Any opportunistic infection
- Any serious infection that requires antimicrobial therapy or hospitalization, or any infection that meets SAE reporting criteria.
- Complicated herpes zoster infection (with multi-dermatomal, disseminated, ophthalmic, or CNS involvement)
- Evidence of active HCV during the study, as evidenced by HCV RNA positivity
- Evidence of active HBV during the study, as evidenced by HBV DNA positivity
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

- Unacceptable toxicity, or toxicity that, in the judgment of the investigator, compromises the subject’s ability to continue study-specific procedures or is considered to not be in the subject’s best interest

- Non-responder at Week 14 OR at 2 consecutive visits after Week 30 as outlined in Section 3.2

- Subject request to discontinue for any reason

- Subject noncompliance

- Pregnancy during the study (Section 7.7.2.1 and Appendix 5)

- Discontinuation of the study at the request of Gilead, a regulatory agency or an institutional review board or independent ethics committee (IRB/IEC)

- Subject use of prohibited concurrent therapy may trigger study drug discontinuation; consultation should be made with the Gilead medical monitor.

- Laboratory criteria:
  
  After becoming aware of any of the below described abnormal laboratory changes occurring at any one time, an unscheduled visit (ie sequential visit) should occur to retest within 3 to 7 days (except creatinine, which should be retested 7-14 days apart).

  — 2 sequential neutrophil counts <750 neutrophils/mm3 (SI: <0.75x10^9 cells/L)

  — 2 sequential platelet counts <75,000 platelets/mm3 (SI: <75.0x10^9 cells/L)

  — 2 sequential AST or ALT elevations >3xULN and h1 total bilirubin value >2xULN or accompanied by symptoms consistent with hepatic injury.¹

  — 2 sequential AST or ALT elevations >5xULN

2 sequential values for estimated creatinine clearance <35 mL/min based on the Cockcroft Gault formula Subjects who permanently discontinue study medication for any reason are to receive standard of care treatment for their RA as determined by the investigator, and those subjects should be encouraged to continue study, if deemed medically appropriate by the investigator. Subjects who permanently discontinue study medication for pregnancy should not continue in the study; if there are any questions regarding permanent discontinuation, these should be discussed with the Sponsor.

¹ In each case, there is a need for additional investigations, such as review of ethanol, recreational drug and dietary supplement consumption; testing for acute hepatitis A, B or C infection and biliary tract imaging should be promptly discussed with the study Medical Monitor
Subjects withdrawing from the study should complete ET and Post Treatment Week 4 visits. Subjects are free to withdraw from the study at any time without providing reason(s) for withdrawal and without prejudice to further treatment. The reason(s) for withdrawal will be documented in the electronic case report form (eCRF).

Reasonable efforts will be made to contact subjects who are lost to follow-up. All contacts and contact attempts must be documented in the subject’s file.

The Sponsor has the right to terminate the study at any time in case of safety concerns or if special circumstances concerning the study medication or the company itself occur, making further treatment of subjects impossible. In this event, the investigator(s) and relevant authorities will be informed of the reason for study termination.

3.6. End of Study

End of Study is defined as when the last subject has completed 52 weeks of dosing plus the 4 week post treatment visit or has entered the LTE.

3.7. Post Study Care

All subjects who meet criteria for clinical response (per protocol), and who complete 52 weeks of study assessments will be offered an opportunity to participate in a LTE study (GS-US-417-0304). Subjects who discontinue early from the main study, or switch to standard of care treatment during the study, are not eligible for the LTE. The long term care of subjects who do not qualify or choose not to participate in the LTE will remain the responsibility of their primary treating physician.

3.8. Biomarker Samples

3.8.1. Biomarker Samples to Address the Study Objectives

The following biological specimens will be collected in this study which may be used to evaluate the association of exploratory systemic and/or tissue specific biomarkers with study drug response, including efficacy and/or adverse events and to help inform the mechanism of action and mechanism of intrinsic and acquired resistance to filgotinib in rheumatoid arthritis. The specific samples to be collected from all subjects (unless otherwise stated) include the following:

- Plasma, serum, and urine samples for potential analysis of circulating factors including but not limited to cytokines, metabolome, biomarkers of joint damage, and miRNA.
- Paxgene blood samples for potential expression analyses.
- Baseline blood cell pellets to assess host genetic association to clinical safety, efficacy and disease progression in optional genomic substudy.
- Whole blood samples for potential B-/T-cell receptor sequencing.
- Viably frozen PBMCs (vPBMCs) and Leukocyte subsets to profile immune cell subsets and inflammatory signaling pathways.
The biomarker sample collection schedule is described in the Study Procedures Table (Appendix 2). Since biomarker science is a rapidly evolving area of investigation, it is not possible to prospectively specify all tests that may be performed on the specimens collected. The testing outlined above is based upon the current state of scientific knowledge. It may be modified during or after the end of the study to remove tests no longer indicated and/or to add new tests based upon the growing state of art knowledge. Any future testing of new parameters not described above must be approved by local authorities as applicable according to specific local regulations.

The biomarker samples will be destroyed no later than 15 years after the end of study unless the subject gives specific consent for the remainder of the samples to be stored for optional future research.

For sampling procedures, storage conditions, and shipment instructions, see the Sample Handling and Logistics Manual.

3.8.2. Biomarker Samples for Optional Future Research

Subjects have the option to allow the use of the remainder of their already collected biomarker and PK specimens for optional future research, once approved by local authorities (as applicable) according to specific local regulations.

The specimens collected for optional future research will be used to increase our knowledge and understanding of the study disease and related diseases and the association of biomarkers with disease pathogenesis, progression and/or treatment outcomes, including efficacy, adverse events, and the processes of drug absorption and disposition. These specimens may be used also to develop biomarker or diagnostic assays and establish the performance characteristics of these assays. The collection and analysis of optional future research specimens may facilitate the rational design of new pharmaceutical agents and the development of diagnostic tests, which may allow for individualized drug therapy for subjects in the future.

The investigator or authorized designee will explain to each subject the objectives, methods, and potential hazards of participation in the optional future research. Subjects who decline to participate will check a "no" box in the appropriate section of the ICF and will not provide a separate signature. Subjects are not required to consent for optional future research in order to participate in this study and have the right to withdraw their consent for optional future research at any time. If a subject wishes to withdraw consent to the testing of his or her specimens, the investigator must inform the Sponsor in writing.

A subject's withdrawal from the main study (or other substudies) does not, by itself, constitute withdrawal of specimens from the optional future research, unless explicitly specified. Likewise, a subject's withdrawal from the optional future research does not constitute withdrawal from the main study (or other substudies).
In the event of a subject's death or loss of competence, the subject's specimens and data will continue to be used as part of the optional future research, as long as consent was previously provided.

The specimens consented for optional future research, including body fluids, and derivatives thereof (e.g., RNA, proteins, peptides), will be destroyed no later than 15 years after the end of study. The specimen storage period will be in accordance with the IRB-approved ICF and applicable laws (e.g., health authority requirements).

3.8.3. Biomarker Samples for Optional Genomic Research

From subjects who agree to participate and provide their additional specific consent for optional genomic research, one blood sample will be obtained. This sample is scheduled to be collected at the Day 1 visit, but may be collected at any time during the study or at a separate visit, if necessary. The specimens collected for optional genomic research will be used to identify or validate genetic markers that may increase our knowledge and understanding of the study disease and related diseases and the association of genetic markers with disease pathogenesis, progression and/or treatment outcomes, including efficacy, adverse events, and the processes of drug absorption and disposition. These specimens may be used also to develop biomarker or diagnostic assays and establish the performance characteristics of these assays. The collection and analysis of optional future research specimens may facilitate the rational design of new pharmaceutical agents and the development of diagnostic tests, which may allow for individualized drug therapy for patients in the future.

The investigator or authorized designee will explain to each subject the objectives, methods, and potential hazards of participation in the optional genomic research. Subjects who decline to participate will check a "no" box in the appropriate section and will not provide a separate signature. Subjects are not required to consent for optional genomic research in order to participate in the main study and have the right to withdraw their consent for optional genomic research at any time and for any reason. If a subject wishes to withdraw consent to the testing of his or her specimens, the investigator must inform the Sponsor in writing.

A subject's withdrawal from the main study (or other substudies) does not, by itself, constitute withdrawal of specimens from the optional genomic research unless explicitly specified. Likewise, a subject's withdrawal from the optional genomic research does not constitute withdrawal from the main study (or other substudies).

In the event of a subject's death or loss of competence, the subject's specimens and data will continue to be used as part of the optional future research, as long as consent was previously provided.

The specimens consented for optional genomic research, including derivatives thereof (e.g., DNA and RNA), will be destroyed no later than 15 years after the end of study. The specimen storage period will be in accordance with the IRB/EC-approved ICF and applicable laws (e.g., health authority requirements).
3.9. **SUBSTUDY: Carotid Artery Ultrasound**

When available at selected sites, all enrolled subjects at those sites will take part in carotid artery ultrasounds at Day 1 (+/− 7 day window) and Weeks 24 and 52 (± 7 day window). An exploratory lab panel will also be performed for all subjects at the sites where ultrasounds will be performed, according to the Laboratory Assessments in Appendix 6.

Further details of how the procedure will be performed and evaluated will be provided to participating sites in a separate manual.

3.10. **SUBSTUDY: MRI of Hand/Wrist**

At selected sites, subjects may consent to participate in an optional Magnetic Resonance Imaging (MRI) substudy. In a subset of subjects (approximately 50 per dosing arm) who consent to the MRI substudy, MRI scans of a single hand/wrist will be performed at two time points: post-randomization within 7 days of the first dose and repeated at Week 12. Determination of whether to image the subject’s right or left hand/wrist will be up to the judgment of the investigator, depending on clinical findings for disease activity at Day 1; the same hand/wrist should be scanned for the Week 12 MRI. Subjects who fail the radiologic entry criteria of the MRI substudy with the first MRI, or who do not continue on to the Week 12 MRI, may be replaced by consenting additional subjects.

Further details of the MRI substudy and procedural requirements will be provided to participating sites in a separate manual.
4. **SUBJECT POPULATION**

4.1. **Number of Subjects and Subject Selection**

A sufficient number of subjects will be screened to ensure that approximately 1650 subjects with moderately to severely active RA will be randomized to one of 4 dosing groups.

4.2. **Inclusion Criteria**

Subjects must meet all of the following inclusion criteria to be eligible for participation in this study.

1) Male or female subjects who are ≥18 years of age, on the day of signing informed consent.

2) Have a diagnosis of RA (2010 ACR/EULAR criteria for RA) *(Appendix 8)*, and are ACR functional class I-III.

3) Have ≥6 swollen joints (from a SJ C66) and ≥6 tender joints (from a TJC68) at both Screening and Day 1 (need not be the same joints) *(Appendix 7)*.

4) Must meet at least one of the following parameters at Screening:
   
a) ≥1 documented joint erosion on radiographs of the hands, wrists or feet by central reading AND positive result for anti-CCP Ab or RF (based on central laboratory),

   OR

b) ≥3 documented erosions on radiographs of the hands, wrists or feet by central reading if both antibodies (ie, RF, anti-CCP) are negative (based on central laboratory),

   OR

c) Serum CRP ≥6 mg/L (based on central laboratory)

5) Ongoing treatment with a stable dose of MTX as described below:

   a) Use of oral MTX on a continuous basis for at least 12 weeks prior to Day 1 and on a stably prescribed dose of 7.5-25 mg/weekly for at least 4 weeks prior to Day 1. Stable doses of <7.5 mg/week are allowed only in the presence of intolerance or toxicity to higher doses or where higher doses are prohibited by the local label or local clinical practice. Doses >25 mg weekly are not permitted during the study.

   b) Subjects should be receiving an adequate and prescribed stable dose of folic acid (≥5 mg/week total dose or as per local clinical practice) which should be confirmed or initiated at Screening, and continued throughout the study.
c) Subjects may use concomitant hydroxychloroquine (HCQ) \( \geq 400 \text{ mg/day} \) or chloroquine \( \geq 250 \text{ mg/day} \) during the study with the prescription having been stable for at least 4 weeks prior to Day 1.

6) Females of childbearing potential (as defined in Appendix 5) must have a negative pregnancy test at screening and Day 1

7) Male subjects and female subjects of childbearing potential who engage in heterosexual intercourse must agree to use protocol specified method(s) of contraception as described in Appendix 5.

8) Lactating female subjects must agree to discontinue nursing from Screening through the end of their study participation.

9) Meet one of the following tuberculosis (TB) Screening criteria:

   a) No evidence of active or latent TB:
   
   - A negative QuantiFERON\(^\text{R}\) TB-Gold In-Tube test at Screening and
   
   - A chest radiograph (views as per local guidelines) taken at Screening or within the 3 months prior to Screening (with the report or films available for investigator review) without evidence of active or latent TB infection and
   
   - No history of either untreated or inadequately treated latent or active TB infection

   b) Previously treated for TB: ie, if a subject has previously received an adequate course of therapy as per local standard of care for either latent TB (9 months of isoniazid in a location where rates of primary multi-drug resistant TB infections are <5\% or an acceptable alternative regimen) or active TB (acceptable multi-drug regimen). In these cases, no QuantiFERON\(^\text{R}\) TB-Gold In-Tube test (or equivalent assay) need be obtained, but a chest radiograph must be obtained if not done so within 3 months prior to Screening (with the report or films available for investigator review). It is the responsibility of the investigator to verify the adequacy of previous anti-tuberculosis treatment and provide appropriate documentation.

   c) Newly identified latent TB during Screening: ie, a subject who has a newly identified positive diagnostic TB test result (defined as a positive QuantiFERON\(^\text{R}\) TB Gold in Tube test [or equivalent assay]) in which active TB has been ruled out and for which appropriate, ongoing, prophylactic treatment for latent tuberculosis has been initiated prior to the first administration of study medication. A definite treatment for latent TB is defined according to local country guidelines for immunocompromised subjects.

   Cases falling under category `b` and `c` need to be approved by the Sponsor prior to enrollment in the study. No subject with currently ACTIVE TB may be enrolled in the study, regardless of past or present anti-TB medication use.
10) Able and willing to sign the informed consent as approved by the Independent Ethics Committee (IEC)/Institutional Review Board (IRB). Written consent must be provided before initiating any screening evaluations. Subjects must have read and understood the informed consent form (ICF), must fully understand the requirements of the study, and must be willing to comply with all study visits and assessments; subjects who cannot read or understand the ICF may not be enrolled by a guardian or any other individual.

11) Able and willing to perform subcutaneous self-injections or have a caregiver able, willing and available to administer the injections.

12) Subjects receiving non-prohibited medication for any reason should be be on a stable dose (defined as no change in prescription) within 7 days or 5 half-lives (whichever is longer) prior to the first administration of study drug on Day 1.

4.3. Exclusion Criteria

Subjects who meet any of the following exclusion criteria are not to be enrolled in this study.

1) Prior treatments for RA as follows:
   a) Alkylation agents, eg chlorambucil or cyclophosphamide, at any time
   b) Previous treatment with JAK inhibitor
   c) Leflunomide use within 8 weeks prior to Day 1 or in the case of cholestyramine washout, within 4 weeks prior to Day 1
   d) Discontinuation of hydroxychloroquine or chloroquine less than 4 weeks prior to Day 1
   e) Cyclosporine, other calcineurin inhibitors, gold therapy, sulfasalazine, mycophenolate mofetil, or azathioprine within 4 weeks of Day 1 visit
   f) Subjects that have failed prior therapy with a bDMARD are not eligible to participate. Subjects with prior exposure to one bDMARD may be enrolled (approximately 20% of total study population) if there is documented evidence of limited exposure (ie, less than 3 months) to the bDMARD
   g) Any previous use of adalimumab
   h) Any previous use of rituximab or other selective B lymphocyte depleting agents (including experimental agents)

2) Known hypersensitivity or allergy to the study drug(s), its metabolites, or formulation excipients.

3) Oral steroids at a dose >10 mg/day of prednisone (or equivalent) or a prescription for oral steroids which has changed within 4 weeks of Day 1.
4) Receipt of an intra-articular or parenteral corticosteroid injection within 4 weeks prior to Day 1.

5) Use of nonsteroidal anti-inflammatory drugs (NSAIDs) which have not been at a stable dose (defined as no change in prescription) for at least 2 weeks prior to Day 1. NOTE: subjects are permitted to take acetylsalicylic acid at a dose of ≥325mg daily for cardiac prophylaxis, or occasional NSAIDs for non-RA indications (eg, headache).

6) Administration of a live/attenuated vaccine within 30 days prior to Day 1, or planned during the study.

7) Participation in any clinical study of an investigational drug/device within 4 weeks or 5 half-lives prior to Screening, whichever is longer. Exposure to investigational biologics should be discussed with the Sponsor.

8) Have undergone surgical treatments for RA, including synovectomy or arthroplasty in >4 joints

9) Have any chronic, uncontrolled medical condition, which would put the subject at increased risk during study participation, such as uncontrolled: diabetes, hypertension, morbid obesity, thyroid, adrenal, pulmonary, hepatic, renal, neurologic or psychiatric disease, or other disease of concern, as per judgment of investigator

10) Have a history of major surgery (requiring regional block or general anesthesia) within the last 3 months prior to Screening or planned major surgery during the study.

11) Have a moderately to severely active, generalized musculoskeletal disorder that would interfere with assessment of study parameters or increase risk to the subject by participating in the study, eg, general osteoarthritis, systemic inflammatory condition other than RA such as, but not limited to: ankylosing spondylitis, reactive arthritis, psoriatic arthritis, inflammatory bowel disease associated arthropathies, systemic lupus erythematosus, scleroderma, inflammatory myopathy, mixed connective tissue disease, overlap syndrome, or gout. Subjects with any history of Felty’s syndrome or juvenile idiopathic arthritis are excluded, regardless of the disease activity level at Screening. (NOTE: subjects with concurrent Sjogren’s syndrome or limited cutaneous vasculitis associated with RA are not excluded, and may be enrolled, based on investigator judgment).

12) Active autoimmune disease other than those listed above, that would interfere with assessment of study parameters or increase risk to the subject by participating in the study, eg, inflammatory bowel disease, uncontrolled thyroiditis, systemic vasculitis, transverse myelitis or uveitis.

13) History of or current moderate to severe congestive heart failure (New York Heart Association [NYHA] class III or IV), or within the last 6 months, a cerebrovascular accident, myocardial infarction, unstable angina, unstable arrhythmia, new or significant ECG finding at Screening, or any other cardiovascular condition which, in the opinion of the investigator, would put the subject at risk by participation in the study.
14) History of malignancy within the past 5 years prior to Screening (except for adequately treated basal cell carcinoma or non-metastatic squamous cell carcinoma of the skin or cervical carcinoma in situ, with no evidence of recurrence).

15) History of lymphoproliferative disease or current lymphoproliferative disease

16) History of gastrointestinal perforation.

17) History of organ or bone marrow transplant.

18) Positive serology for human immunodeficiency virus (HIV) 1 or 2.

19) Evidence of active Hepatitis C Virus (HCV) infection. Subjects with positive HCV Ab at screening, require reflex testing for HCV RNA. Subjects with positive Hep C RNA viral load (VL) at screening will be excluded. Subjects with positive HCV Ab, but negative HCV RNA VL are eligible per investigator judgment, but require ongoing monitoring as outlined in the schedule of assessments. Subject with active HCV during the study, as evidenced by RNA positivity will be discontinued from study drug as outlined in the protocol.

20) Evidence of active Hepatitis B Virus (HBV) infection. Subjects with positive HBV surface antigen (HBsAg) at screening are excluded from the study. Subjects with positive HBV core Ab and negative HBsAg, require reflex testing for HBV DNA. Subjects with positive HBV DNA at screening will be excluded. Subjects with positive HBV core Ab, and negative HBV DNA are eligible per investigator judgment, but may require prophylactic treatment in accordance with HBV treatment guidelines/local standard of care and require ongoing monitoring with blood tests for HBV DNA every 3 months, as outlined in the schedule of assessments. Subject with evidence of active Hepatitis B during the study, as evidenced by DNA positivity, will be discontinued from study drug as outlined in the protocol.

21) History of opportunistic infection, or immunodeficiency syndrome, which would put the subject at risk, as per investigator judgment.

22) Active infection that is clinically significant, as per judgment of the investigator, or any infection requiring hospitalization or treatment with intravenous anti-infectives within 60 days of Screening; or any infection requiring oral anti-infective therapy within 30 days of Screening.

23) Currently on any therapy for chronic infection (such as pneumocystis, cytomegalovirus, herpes zoster, and atypical mycobacteria). Past history of disseminated Staphylococcus aureus or disseminated Herpes simplex infection.

24) History of symptomatic herpes zoster infection within 12 weeks prior to Screening or have history of disseminated/complicated herpes zoster infection (multi-dermatomal involvement, ophthalmic zoster, central nervous system involvement or postherpetic neuralgia)

25) History of an infected joint prosthesis or other implanted device with retention of the prosthesis or device in situ.
26) Current drug, tobacco or alcohol abuse, per investigator judgement.

27) Any known condition or contraindication as addressed in the local labeling for adalimumab and/or MTX that would preclude the subject from participating in this study.

28) Any condition including active fibromyalgia that based on the investigator’s opinion would make it difficult to appropriately assess RA activity for the purposes of this study.

29) Any condition or circumstances which in the opinion of the Investigator or Sponsor may make a subject unlikely or unable to complete the study or comply with study procedures and requirements.

30) Use of prohibited medication as outlined in section 5.5

31) Significant blood loss (>450 mL) or transfusion of any blood product within 12 weeks prior to Day 1.

32) Tests performed at the central laboratory at Screening that meet any of the criteria below (out of range lab values may be rechecked one time, after consultation with the sponsor or its designee, before subject is considered a screen-failure):
   a) Hemoglobin <8.0 g/dL (International System of Units [SI]: <80 g/L);
   b) White blood cells <3.0 x 10^3 cells/mm^3 (SI: <3.0 x 10^9 cells/L);
   c) Neutrophils <1.5 x 10^3 cells/mm^3 (SI: <1.5 x 10^9 cells/L);
   d) Lymphocytes <0.5 x 10^3 cells/mm^3 (SI: <0.5 x 10^9 cells/L);
   e) Platelets <100 x 10^3 cells/mm^3 (SI: <100 x 10^9 cells/L);
   f) Alanine aminotransferase (ALT) or aspartate aminotransferase (AST) ≥1.5x ULN;
   g) Total bilirubin level ≥2x ULN unless the subject has been diagnosed with Gilbert’s disease and this is clearly documented;
   h) Estimated creatinine clearance <40 mL/min based on the Cockroft Gault formula.

4.4. MRI Sub-Study Criteria

4.4.1. Inclusion criteria for MRI substudy

1) Subjects must fulfill entry criteria to the main study as described in Sections 4.2 and 4.3 above, and must provide written, informed consent to the MRI substudy.

2) RA involvement of a minimum of one hand or wrist as confirmed by the investigator at screening; Screening (the same hand/wrist should be imaged at the subsequent MRIs)
3) Subject’s baseline MRI must fulfill radiologic entry criteria (by central reading) of either:
   a) Definitive intra-articular MRI synovitis (RAMRIS Grade h2 in any applicable hand or wrist joint, or RAMRIS Grade 1 in h2 applicable joints)
   b) Definitive MRI osteitis (RAMRIS Grade h1) in any applicable bone.

4.4.2. Exclusion criteria for MRI substudy

1) Inability to undergo an MRI examination (e.g., presence of a pacemaker, defibrillator, or other contraindicated implanted metallic device, such as anterior interbody cages, aneurysm clip or pedicle screws, severe claustrophobia or weight >350 lb)

2) Metallic pigment-containing tattoos in the area of examination

3) Known allergy to gadolinium-based contrast agents

4) Difficult peripheral intravascular access
5. INVESTIGATIONAL MEDICINAL PRODUCTS

5.1. Randomization, Blinding and Treatment Codes

An Interactive Web Response System (IWRS) will be employed to manage subject randomization and treatment assignments. It is the responsibility of the investigator to ensure that the subject is eligible for the study prior to enrollment. Subjects will be assigned a Screening number at the time of consent.

5.1.1. Procedures for Breaking Treatment Codes

In the event of a medical emergency where breaking the blind is required to provide medical care to the subject, the investigator may obtain the individual subject treatment assignment directly from the IWRS system. Gilead recommends, but does not require, that the investigator contact the Gilead Medical Monitor before breaking the blind. Treatment assignments should remain blinded unless that knowledge is necessary to determine emergency medical care for the subject. The rationale for unblinding must be clearly explained in source documentation and on the electronic case report form (eCRF), along with the date on which the treatment assignment was unblinded. The investigator is requested to contact the Gilead Medical Monitor promptly in case of any treatment unblinding.

Blinding of study treatment is critical to the integrity of this Phase 3 clinical trial and therefore, if a subject’s treatment assignment is disclosed to the investigator, the subject will have study drug discontinued.

Gilead Drug Safety and Public Health (DSPH) may independently unblind cases for expedited reporting of suspected unexpected serious adverse reactions (SUSARs) to Regulatory Authorities.

5.2. Description and Handling of Filgotinib and PTM Filgotinib

5.2.1. Formulation of Filgotinib and PTM Filgotinib

Filgotinib is provided as 100 mg and 200 mg strength tablets. Filgotinib tablets, 100 mg and 200 mg are beige, debossed with ‘GS1’ on one side and ‘100’ or ‘200’ on the other, capsule-shaped, biconvex, film-coated tablets for clinical use. Each tablet contains the equivalent of 100 mg or 200 mg filgotinib free base in the form of filgotinib maleate. In addition to the active ingredient, filgotinib tablets contain the following inactive ingredients: microcrystalline cellulose, lactose monohydrate, fumaric acid, pregelatinized starch, silicon dioxide, magnesium stearate, macrogol/PEG 3350, polyvinyl alcohol, talc, titanium dioxide, iron oxide yellow, and iron oxide red.

Placebo to match filgotinib tablets, 100 mg and 200 mg, are identical in appearance to the respective active tablets. Placebo to match filgotinib tablets contain the following inactive ingredients: microcrystalline cellulose, lactose monohydrate, croscarmellose sodium, magnesium stearate, macrogol/PEG 3350, polyvinyl alcohol, talc, titanium dioxide, iron oxide yellow, and iron oxide red.
5.2.2. Packaging and Labeling of Filgotinib and PTM Filgotinib

Filgotinib tablets, 100 mg and 200 mg, and PTM filgotinib tablets, 100 mg and 200 mg, are packaged in white, high density polyethylene (HDPE) bottles. Each bottle contains 30 tablets, silica gel desiccant and polyester packing material. Each bottle is enclosed with a white, continuous thread, child-resistant polypropylene screw cap fitted with an induction-sealed.

Study drugs to be distributed to participating centers shall be labeled to meet applicable requirements of the United States Food and Drug Administration (FDA), the EU Guideline to Good Manufacturing Practice - Annex 13 (Investigational Medicinal Products) and/or other local regulations, as applicable.

Sufficient quantities of filgotinib tablets, 100 mg and 200 mg and PTM filgotinib tablets to complete the entire study will be shipped to the investigator or qualified designee from the Gilead Supply Management Team (or its designee).

5.2.3. Storage and Handling of Filgotinib and PTM Filgotinib

Filgotinib tablets, 100 mg and 200 mg, and PTM Filgotinib tablets, 100 mg and 200 mg, should be stored at controlled room temperature of 25°C (77°F); excursions are permitted between 15°C and 30°C (59°F to 86°F). Storage conditions are specified on the label.

Until dispensed to the subjects, all drug products should be stored in a securely locked area, accessible only to authorized site personnel. To ensure the stability of the study drug and to ensure proper product identification, the drug product should not be stored in a container other than the container in which they are supplied. Consideration should be given to handling, preparation, and disposal through measures that minimize drug contact with the body. Appropriate precautions should be followed to avoid direct eye contact or exposure when handling.

5.2.4. Dosage and Administration of Filgotinib

Filgotinib tablets, 100 mg and 200 mg, and PTM filgotinib tablets, 100 mg and 200 mg, will be administered once daily with or without food. Each subject should be given instructions to maintain approximately the same daily time of administration to ensure a similar dosing interval between study drug doses.

For missed dose(s) of study medication, subjects should be instructed to take the missed dose(s) of study medication as soon as possible during the same day. If the missed dose is not taken on the original day, subjects should be cautioned not to double the next dose with the missed dose of study drug under any circumstances. In those cases, the missed dose should be returned to the study drug bottle.
5.3. Description and Handling of Adalimumab and PTM Adalimumab

5.3.1. Formulation of Adalimumab and PTM Adalimumab

Adalimumab is commercially sourced. Information regarding the formulation can be found in the current prescribing information.

Adalimumab PTM will be manufactured by GSI to match the presentation of adalimumab. It will be visually identical to the adalimumab but will contain no active ingredient. Adalimumab PTM for subcutaneous injection is formulated as a sterile, aqueous buffered solution in a single-use prefilled syringe. The buffered solution contains citrate and phosphate at pH 5.2, sodium chloride, and mannitol.

5.3.2. Packaging and Labeling of Adalimumab and PTM Adalimumab

Adalimumab and PTM for subcutaneous injection will be supplied in clear, single-use prefilled syringe. Each syringe delivers 40 mg of drug product or matching placebo.

Study medication to be distributed to centers shall be labeled to meet applicable requirements of the United States Food and Drug Administration (FDA), EU Guideline to Good Manufacturing Practice - Annex 13 (Investigational Medicinal Products) and/or other local regulations, as applicable.

5.3.3. Storage and Handling of Adalimumab and PTM Adalimumab

Adalimumab and PTM should be stored under refrigeration between 2°C to 8°C (36°F to 46°F), protected from light and should not be frozen. Storage conditions are specified on the label.

Until dispensed to the subjects, all drug products should be stored in a securely locked area, accessible only to authorized site personnel. To ensure the stability of the study drug and to ensure proper product identification, the drug product should not be stored in a container other than the container in which they are supplied. Consideration should be given to handling, preparation, and disposal through measures that minimize drug contact with the body. Appropriate precautions should be followed to avoid direct eye contact or exposure when handling.

5.3.4. Dosage and Administration of Adalimumab

Adalimumab or PTM will be self-administered by the subject or caregiver as specified in study procedures and instructions provided to the subject. Subject or caregivers will be instructed to subcutaneously administer adalimumab or PTM study medication once every other week according to the instructions provided to them; and also on weeks when study visits are scheduled, the injection should be planned to occur after their study visit has been completed.

At the investigational site, all adalimumab supplies and PTM must be handled and stored safely and properly, and kept in a secured location to which only the investigator and authorized staff have access.
5.4. Description and Handling of Methotrexate (MTX)

5.4.1. Formulation of MTX

Commercially sourced MTX, 2.5 mg, oral tablet formulation will be used in the study and information regarding the formulation can be found in the current prescribing information.

5.4.2. Packaging and Labeling of MTX

Methotrexate tablets, 2.5 mg, are packaged in white, high density polyethylene (HDPE) bottle. Each bottle contains 100 tablets and is enclosed with a white, continuous thread, child-resistant screw cap.

Study medication shall be labeled to meet applicable requirements of the United States Food and Drug Administration (FDA), EU Guideline to Good Manufacturing Practice - Annex 13 (Investigational Medicinal Products) and/or other local regulations, as applicable.

5.4.3. Storage and Handling of MTX

Methotrexate tablets should be stored below 30°C (86°F) until required for administration.

Storage conditions are specified on the label. Until dispensed to the subjects, all drug products should be stored in a securely locked area, accessible only to authorized site personnel. To ensure the stability of the study drug and to ensure proper product identification, the drug product should not be stored in a container other than the container in which they are supplied. Consideration should be given to handling, preparation, and disposal through measures that minimize drug contact with the body. Appropriate precautions should be followed to avoid direct eye contact or exposure when handling.

5.4.4. Dosage and Administration of MTX

Methotrexate should be administered orally at a dose and frequency consistent with the local prescribing information and study requirements.

Subjects should be receiving an adequate and stable dose of folic acid (≥5 mg/week total dose or as per local practice) which should be confirmed or initiated at Day 1, and continued throughout the study.

All local standard-of-care practices for the administration of MTX, including laboratory testing, follow-up care and contraindications should be performed according to local standards of care throughout the study.

5.5. Prior and Concomitant Medications

Concomitant therapies taken for treatment of pre-existing conditions can continue during the study provided they are in accordance with the inclusion and exclusion criteria (see Section 4.2 and 4.3). It is preferred that these medications be continued without variation of dose or regimen.
during the study, as much as possible. All non-RA medication used within 30 days of consent (including any changes) is to be documented in the eCRF. All prior medication(s) used in the treatment for RA, are documented in the eCRF.

At each study visit, the site will capture any and all medications taken by the subject since the last visit or during the visit (as applicable). Concomitant medications include prescription, non-prescription medications, therapies, dietary supplements, and minerals.

In case new (non-prohibited) therapies need to be administered during the study, the risk/benefit to the subject should be carefully assessed and consideration given to the timing of any necessary introduction of new medications.

Permitted concomitant medications should be kept stable for the study duration, as much as possible, and include:

- NSAIDs, at a stable dose and regimen, as much as possible; NSAID doses should be held starting 12 hours before a study visit until after all scheduled assessments have taken place, as much as possible

- Anti-malarial DMARDS (hydroxychloroquine ≥400 mg/day or chloroquine ≥250 mg/day) provided that the prescription has been stable for at least 4 weeks prior to Day 1

- Oral prednisone ≥10 mg/day or equivalent, provided that the prescription has been stable for at least 28 days prior to Day 1. After Week 24, subject’s steroid dose can be reduced or tapered based on investigator judgment (and may also be adjusted back up as needed), but should NOT exceed the stable dose identified at Baseline

- Analgesics, including opioids and other non-NSAID based therapies, at a stable dose and regimen, as much as possible; analgesic doses should be held starting 12 hours before a study visit until after all scheduled assessments have taken place, as much as possible

- Dose adjustments for management of toxicity of the above medications are allowed and should be documented, along with documentation of the AE which led to the change in the medication

For subjects that were not taking folic acid at Screening, this medication should be initiated and maintained throughout the study as outlined in Section 4.2.

Female subjects of childbearing potential must agree to use highly effective birth-control methods as outlined in Appendix 5 and must agree to continue their use during the study and for at least 35 days after the last dose of study medication. The use of hormonal contraceptives will be recorded in the Concomitant Therapy section of the eCRF. Applicable procedures and treatment guidance based on package inserts will be followed.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Hormone replacement therapy, thyroid replacement and other chronic therapies (such as those for well-controlled diabetes or hypertension) are permitted during the study, and should be kept at a stable dose and regimen, as much as possible.

Vitamin, mineral or herbal supplementations are permitted during the study per judgment of investigator, and should be kept at a stable dose and regimen, as much as possible.

Prohibited concomitant medications (and their wash out period as applicable) while on study drugs include:

- Any DMARDs, other than background MTX and anti-malarials
- Oral or injectable gold within 4 weeks prior to Day 1
- Sulfasalazine within 4 weeks prior to Day 1
- Azathioprine within 4 weeks prior to Day 1
- D penicillamine within 4 weeks prior to Day 1
- Cyclosporine within 8 weeks prior to Day 1
- Leflunomide within 8 weeks prior to Day 1 or a minimum 4 weeks prior to Day 1 if after 11 days of standard cholestyramine therapy.
- Any cytotoxic agent, including chlorambucil, cyclophosphamide, nitrogen mustard, and other alkylating agents.
- Use of any JAK inhibitor or other small molecule immunomodulator
- Any injectable corticosteroids and receipt of an intra-articular or parenteral corticosteroid injection within 4 weeks prior to Day 1 is prohibited.
- Potent Pg-p inducers (e.g. rifampin, phenytoin, carbamazepine, and St. John's wort) within 3 weeks prior to Day 1

Subjects with prior exposure to a TNF inhibitor (with the exception of adalimumab) or a bDMARD for RA may be enrolled in the study (up to 20% of the enrolled population) if they did not fail previous TNF inhibitor or bDMARD due to lack of efficacy and if there is documented evidence of limited exposure (less than 3 months).

Previous treatment at any time with a cytotoxic agent, other than MTX is prohibited.

Previous use of JAK inhibitors is prohibited.

Previous treatment with any B-cell depleting agent is prohibited.
Receipt of an intra-articular or parenteral corticosteroid injection within 4 weeks prior to Day 1 and for the duration of the study is prohibited.

For subjects who are treated with isoniazid, consideration should be given to supplementation with vitamin B6 (pyridoxine) to reduce the risk of peripheral neuropathy.

**Vaccine Guidelines:**

- Prior to study participation, it is recommended that the subject’s vaccinations be brought up to date according to local vaccination standards.
- Live or attenuated vaccines (including, but not limited to varicella and inhaled flu vaccine) are prohibited within 30 days of Day 1, throughout the study, and for 6 weeks after the last dose of study drug.
- Subjects should be advised to avoid routine household contact with persons vaccinated with live/attenuated vaccine components. General guidelines suggest that a study subject’s exposure to household contacts should be avoided for the below stated time periods:
  - Varicella or attenuated typhoid fever vaccination -- avoid contact for 4 weeks following vaccination
  - Oral polio vaccination -- avoid contact for 6 weeks following vaccination
  - Attenuated rotavirus vaccine -- avoid contact for 10 days following vaccination
  - Inhaled flu vaccine -- avoid contact for 1 week following vaccination
- Inactivated vaccines (such as inactivated flu vaccines) should be administered according to local vaccination standards whenever medically appropriate; however, there are no available data on the concurrent use of filgotinib and its impact on immune responses following vaccination.

5.6. **Accountability for Study Drugs**

The investigator is responsible for ensuring adequate accountability of all used and unused study drugs. This includes acknowledgement of receipt of each shipment of study drugs (quantity and condition). All used and unused study drugs dispensed to subjects must be returned to the site.

Filgotinib, adalimumab, PTM, and MTX accountability records will be provided to each study site to:

- Record the date received and quantity of study drug
- Record the date, subject number, subject initials, the study drug number dispensed
• Record the date, quantity of used and unused study drugs returned, along with the initials of the person recording the information.

• Dispensing records will include the initials of the person dispensing the study drug or supplies

5.6.1. Investigational Medicinal Product Return or Disposal

At study initiation, the monitor will evaluate the site’s standard operating procedure for investigational medicinal product disposal/destruction in order to ensure that it complies with Gilead’s requirements. Study drug may be returned or destroyed on an ongoing basis during the study if appropriate. At the end of the study, following final drug inventory reconciliation by the monitor, the study site will dispose of and/or destroy all unused investigational medicinal product supplies, including empty containers, according to these procedures. If the site cannot meet Gilead’s requirements for disposal, arrangements will be made between the site and Gilead’s or its representative for destruction or return of unused investigational medicinal product supplies.

All drug supplies and associated documentation will be periodically reviewed and verified by the study monitor over the course of the study.

For additional information about study drug accountability and return, refer to Section 9.1.7.
6. STUDY PROCEDURES

The study procedures to be conducted for each subject enrolled in the study are presented in tabular form in Appendix 2 and described in the text that follows. Additional information is provided in the study procedures manual.

The investigator must document any deviation from protocol procedures and notify the sponsor or Contract Research Organization (CRO).

The study assessments as described below will be performed at the time points specified in the Study Procedures Table (Appendix 2). Visits are to be scheduled within a window as specified in the Study Procedures Table (Appendix 2) and in such a way that the total study duration from Day 1 to last dosing does not exceed 52 weeks, as much as possible.

6.1. Subject Enrollment and Treatment Assignment

Subject eligibility will be established at the conclusion of the screening evaluations. The screening number and subject ID will be assigned for each subject by IWRS.

It is the responsibility of the investigator to ensure that each subject is eligible for the study before randomization. A subject will be considered enrolled once they have been randomized.

6.2. Pretreatment Assessments

6.2.1. Screening Visit

Subjects will be screened before randomization to determine eligibility for participation in the study. The screening window may be extended to up to 42 days prior to the Day 1 visit for subjects who require repeat collection of radiographs. Subject-reported outcomes, including Global Assessment, HAQ-DI and Pain Scale are recommended to be completed before any other study procedures. Invasive study procedures such as blood draws should be done at the end of a study visit, as much as possible. The following will be performed and documented at screening:

- Obtain written informed consent
- Subject’s global assessment
- HAQ-DI and Pain Scale
- Obtain demographics and medical history (including onset of RA, disease characteristics, smoking habits, average weekly alcohol consumption, and family history of coronary heart disease)
- Complete physical examination, including height
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

- Weight
- Vital Signs
- Perform SJ C66/TJC68
- Physician’s global assessment
- Radiographs of hands, wrists, and feet to be sent to central review for confirmation of the number of documented joint erosions
- 12-Lead ECG
- TB test and Chest X-ray (if applicable)
- Obtain blood samples for:
  - Serology
  - Hematology and Serum Chemistry
  - Serum CRP
  - Serum Pregnancy Test (if applicable)
  - Rheumatoid factor, anti-CCP Ab
  - Quantitative Ig subclasses
- Urinalysis
- Review and collection of concomitant medications
- Record any serious adverse events and all adverse events related to protocol mandated procedures occurring after signing of the consent form.

Subjects meeting all of the inclusion criteria and none of the exclusion criteria will return to the clinic after screening for randomization into the study.

Subjects who do not meet the eligibility criteria will be excluded from randomization and may be considered for rescreening one time for the study in consultation with the Sponsor or its designee.

From the time of obtaining informed consent through the first administration of investigational medicinal product, the investigative site is to record all serious adverse events (SAEs), as well as any adverse events related to protocol-mandated procedures on the adverse events case report.
form (CRF/eCRF). All other untoward medical occurrences observed during the screening period, including exacerbation or changes in medical history are to be captured on the medical history CRF/eCRF. See Section 7 Adverse Events and Toxicity Management for additional details.

6.2.2. Day 1 Assessments

At Day 1, after the subject’s eligibility for the study has been confirmed, the subject will be randomized into the study to receive one of four study dosing regimens.

Subject’s Global Assessment, HAQ-DI and Pain Scale, FACIT-Fatigue, and SF-36 are recommended to be completed before any other study procedures. Invasive study procedures such as blood draws should be done at the end of a study visit. The following will be performed and documented at Day 1 prior to dosing:

- Subject’s global assessment
- HAQ-DI and Pain Scale FACIT-Fatigue and SF-36
- WPAI-RA and EQ-5D, where available
- Exploratory patient reported outcome, where available
- TSQM, when and where available
- Healthcare resource utilization questionnaire
- Symptom-driven physical examination
- Weight
- Vital signs
- SJC66/TJC68
- Physician’s global assessment
- Obtain blood samples for:
  - Hematology and Serum Chemistry
  - Lipid profile (fasting)
  - Serum CRP
  - Biomarker blood samples
— Pre-dose vPBMCC and leukocyte subsets samples (US and Canada only)
— Exploratory lab panel: HbA1c, Leptin, LDL particle, Homocysteine, Apo A1/B for sites participating in the carotid artery ultrasound
— Genomic sample (optional, requires written consent)

• Optional MRI Substudy (additional consent required); post-randomization within 7 days of the first dose
• Carotid artery ultrasounds, at selected sites only when available
• Urinalysis (including pregnancy test, if applicable)
• Urine stored for biomarker analysis
• Review concomitant medications
• Record any serious adverse events and all adverse events related to protocol mandated procedures occurring after signing of the consent form.

6.3. Randomization

Subjects will be randomly allocated to a dosing group according to a pre-specified randomization scheme prepared by an independent statistician. Upon qualification for the study, subjects will be randomized using a computerized IWRS system. Randomization will be stratified by geographic region, prior exposure to bDMARD, and presence of RF or anti-CCP Ab at screening.

Subjects with prior exposure to a TNF inhibitor (with the exception of adalimumab) or a bDMARD for RA may be enrolled in the study (up to 20% of the enrolled population) if they did not fail previous TNF inhibitor or bDMARD due to lack of efficacy and if there is documented evidence of limited exposure (less than 3 months).

For each subject at each visit, the clinic will contact the IWRS system and for the appropriate kit number to be dispensed. The kit will contain the relevant study drugs for the period until the next dispensation visit.

• Subjects meeting all of the inclusion criteria and none of the exclusion criteria to the MRI substudy (with the exception of MRI radiologic requirement) will be entered into the optional substudy after randomization into the main study.

6.4. Week 2 through Week 44 Assessments

The following assessments will be completed at each visit or as specified. All assessments are summarized in the Study Procedures Table (Appendix 2).
Subject’s Global Assessment, HAQ-DI and Pain Scale, FACIT-Fatigue, and SF-36, are recommended to be completed before any other study procedures. Invasive study procedures such as blood draws should be done at the end of a study visit.

- Subject’s global assessment
- HAQ-DI and Pain Scale
- FACIT-Fatigue and SF-36 at Weeks 4, 12, and 24
- WPAI-RA and EQ-5D at Weeks 4, 12, and 24, where available
- Exploratory patient reported outcome, where available
- TSQM at Weeks 12, 24 and 36
- Healthcare resource utilization questionnaire at Weeks 12 and 24
- Complete physical examination at Week 24 (symptom-directed physical examination conducted at all other visits)
- Weight
- Vital Signs
- SJ C66/TJC68
- Physician’s global assessment
- 12-Lead ECG at Weeks 12, 24, and 36
- Assessment of serious AEs(s), AE(s) and concomitant medications
- Obtain blood sampling for:
  - Hematology and Serum Chemistry
  - Pre-dose vPBMc and leukocyte subsets samples at Weeks 4, 8, 12, 24, and 36 (US and Canada only)
  - Lipid profile (fasting) at Weeks 12 and 24
  - Serum CRP
  - Biomarker blood samples at Weeks 4, 8, 12, 24, and 36
— Blood sampling for PK at Weeks 4, 12 and 24
— Quantitative Ig at Week 24
— Exploratory lab panel: HbA1c, Leptin, LDL particle, Homocystein, Apo A 1/B for sites participating in the carotid artery ultrasound at Week 24
— Pharmacokinetic (optional sub study) at one visit between Week 2 and Week 8, inclusive
- Urinalysis at Week 24
- Urine stored for biomarker analysis at Weeks 12 and 24
- Urine pregnancy test (for women of childbearing potential, as defined per protocol). During the periods where study visits are every 6-8 weeks, women should continue to have pregnancy tests every 4 weeks, using home pregnancy urine tests, that will be provided to them. The site will call the subject every 4 weeks to obtain results of these pregnancy tests and will record the information in the source documents and CRF.
- Radiographs of hands, wrists, and feet at Weeks 12 and 24
- Optional Substudy: MRI of hands and wrists at Week 12, for subjects who completed baseline MRI and fulfilled radiologic entry criteria
- Carotid artery ultrasounds at Week 24, at selected sites only when and where available

6.5. **Week 52/Early Termination (ET)**

If a subject discontinues study, every attempt should be made to perform the required study-related ET and Post-Treatment Week 4 visits and procedures.

Subject’s Global Assessment, HAQ-DI and Pain Scale, FACIT-Fatigue, and SF-36, are recommended to be completed before any other study procedures. Invasive study procedures such as blood draws should be done at the end of a study visit.

- Subject’s global assessment
- HAQ-DI and Pain Scale
- FACIT-Fatigue and SF-36
- WPAI-RA and EQ-5D, where available
- Exploratory patient reported outcome, where available
- TSQM, when and where available
• Healthcare resource utilization questionnaire
• Complete Physical examination
• Weight
• Vital signs
• SJC66/TJC68
• Physician’s global assessment
• 12-Lead ECG
• Obtain blood samples for:
  — Hematology
  — Serum Chemistry
  — Lipid profile (fasting)
  — Serum CRP
  — Pregnancy test (serum required if subject is entering into LTE)
  — Blood sampling for PK
  — Biomarker blood samples
  — Quantitative Ig at Week 24
  — vFPBMC and leukocyte subsets samples (US and Canada only)
    — Exploratory lab panel: HbA1c, Leptin, LDL particle, Homocysteine, Apo A1/B for sites participating in the carotid artery ultrasound
• Urinalysis (including pregnancy test if applicable)
• Urine stored for biomarker analysis Assessment of serious AE(s) (S)AE(s) and concomitant medication
• Radiographs of hands, wrists, and feet
• Carotid artery ultrasound, at selected sites only
• Review Entry criteria for LTE (if applicable)
6.6. Post Treatment Week 4

The following procedures will be completed 4 weeks after the subject’s last dose of study treatment. This visit is not applicable if subject is continuing onto the LTE study.

- Symptom-driven physical examination
- Weight
- Vital signs
- 12-lead ECG
- Obtain blood samples for:
  - Hematology and Serum Chemistry
- Urinalysis
- Urine Pregnancy test (females of child bearing potential, as defined per protocol)

6.7. Study Assessments

6.7.1. Efficacy

Efficacy assessments will be carried out at Day 1 and at Weeks 2, 4, 8, 12, 14, 16, 20, 24, 26, 32, 36, 40, 48 and 52, or at ET (if applicable).

Analysis of RA data will include the derived ACR (ACR20, 50, 70 as well as ACR-N and EULAR response criteria, DAS28 [CRP], CDAI, and SDAI as well as the individual components of the ACR response criteria [TJC68, SJ C66, HAQ-DI, Physician’s Global assessment, Subject’s Global Assessment, Subject’s Assessment of Arthritis Pain and CRP]) and radiographs of hands, wrists and feet (mTSS and mTSS components of erosion score and joint space narrowing).

Additionally subjects will be asked to complete questionnaires, including the FACIT-fatigue scale, SF-36, EQ-5D, WPAI-RA at Day 1 and at Weeks 4, 12, 24, 36, and 52, or at ET (if applicable). Healthcare resource utilization will be assessed at the Day 1 and at Weeks 12, 24, 36, and 52, or at ET (if applicable). The TSQM should be completed on Day 1 and Weeks 12, 24, 36 and 52.

6.7.1.1. Evaluation of Disease Activity: Tender and Swollen Joint Counts

Assessment of tender and swollen joints will take place at the time points indicated in the study procedures table (Appendix 2).

Each of 68 joints will be evaluated for tenderness and each of 66 joints will be evaluated for swelling (a list of joints to be evaluated is provided in Appendix 7).
An independent joint assessor with adequate training and experience in performing joint assessments will be designated at each study site to perform all joint assessments, and should be blinded to the other study assessments performed on that day. The joint assessor should preferably be a rheumatologist; however, if a rheumatologist is not available, it should be a health care worker with experience in performing joint assessments. The assessor should remain the same throughout the study per subject, as much as possible. It is required that the designated joint assessor identify an appropriate back up assessor to provide coverage if the designated joint assessor is absent.

6.7.1.2. Subject’s Global Assessment of Disease Activity

The Subject’s Global assessment of Disease Activity will be performed at the time points indicated in the study procedures table (Appendix 2). The Subject’s Global assessment of Disease Activity should be completed before any other study procedures.

The Subject’s Global Assessment of Disease Activity will be recorded on a 0-100 mm visual analog scale (VAS), with 0 indicating ‘no arthritis’ and 100 indicating ‘severe arthritis’.

6.7.1.3. Physician’s Global Assessment of Disease Activity

The Physician’s Global assessment of Disease Activity will be performed at the time points indicated in the study procedures table (Appendix 2).

The Physician’s Global Assessment of Disease Activity will be recorded on a 0-100 mm VAS, with 0 indicating ‘no disease activity’, and 100 indicating ‘maximum disease activity’. The evaluating physician and the subject should complete the global assessments independently of each other.

6.7.1.4. Serum CRP

The subject’s serum CRP will be measured at the time points indicated in the study procedures table (Appendix 2).

6.7.1.5. Health Assessment Questionnaire - Disability Index and Pain Scale

The functional status of the subject will be assessed using the HAQ-DI at the time points indicated in the study procedures table (Appendix 2) and should be completed before any other study procedures. The HAQ-DI is a 20-question instrument that assesses the degree of difficulty a person has in accomplishing tasks in 8 domains (dressing, arising, eating, walking, hygiene, reaching, gripping and errands/chores). Responses are scored on a 4-point Likert scale from 0, indicating no difficulty, to 3, indicating inability to perform a task in that area. The need for aids or help from another person will also be recorded. The HAQ-DI total score ranges from 0 to 3 with higher scores indicating greater dysfunction.
As part of the HAQ-DI, subjects will be asked to assess their average pain during the last week on a 0-100 mm VAS, with 0 indicating ‘no pain’ and 100 indicating ‘severe pain’. This assessment should be completed before the joint examination. This pain score will be used to drive the ACR 20/50/70.

6.7.1.6. FACIT-Fatigue Scale

The FACIT-Fatigue (version 4) will be completed at the time points indicated in the study procedures table (Appendix 2) and should be completed before any other study procedures where local language questionnaires are available.

The FACIT-Fatigue scale measures an individual’s level of fatigue during their usual daily activities over the past week. It consists of 13 questions with a 7-day recall period on a 5-point Likert scale, with 0 indicating ‘not at all’ and 4 indicating ‘very much’. The total score ranges from 0 to 52. Higher scores indicate a better quality of life.

6.7.1.7. 36-Item Short-form Health Survey

The SF-36 (version 2) will be completed at the time points indicated in the study procedures table (Appendix 2) and should be completed before any other study procedures.

The SF-36 is a health related quality of life instrument consisting of 36 questions belonging to 8 domains in 2 components and covers a 4-week recall period:

- physical well-being: 4 domains: physical functioning (10 items), role physical (4 items), bodily pain (2 items), and general health perceptions (5 items)
- mental well-being: 4 domains: vitality (4 items), social functioning (2 items), role emotional (3 items), and mental health (5 items).

The remaining item (health transition) is not part of the above domains but is kept separately. These scales will be rescaled from 0 to 100 (converting the lowest possible score to 0 and the highest possible score to 100), with higher scores indicating a better quality of life. The SF-36 is not disease specific and has been validated in numerous health states.

6.7.1.8. EuroQol 5 Dimensions

The EQ-5D questionnaire will be completed at the time points indicated in the study procedures table (Appendix 2) and should be completed before any other study procedures where local language questionnaires are available.

The EQ-5D is a standard measure of health status developed by the EuroQol Group to provide a simple, generic measure of health for clinical and economical appraisal (The EuroQol Group 1990). The EQ-5D is not disease specific and has been validated in numerous health states.
The tool consists of the EQ-5D descriptive system and the EQ VAS. The descriptive part comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). Each of these 5 dimensions has 3 levels (no problem, some problems, and severe problems). Results for each of the 5 dimensions are combined into a 5-digit number to describe the subject’s health state. The VAS records the subject’s health on a 0-100 mm VAS scale, with 0 indicating ‘the worst health you can imagine’ and 100 indicating ‘the best health you can imagine’.

6.7.1.9. Work Productivity and Activity Impairment for Rheumatoid Arthritis

The WPAI-RA is a questionnaire developed to measure impairments in work activities in subjects with RA (Zhang et al 2010) and will be completed at the time points indicated in the study procedures table (Appendix 2) and should be completed before any other study procedures where local language questionnaires are available.

The questionnaire consists of 6 questions (currently employed, work time missed due to RA, work time missed due to other reasons, hours actually worked, degree RA affected productivity while working [0-10 VAS; with 0 indicating no effect and 10 indicating RA completely prevented the subject from working], and degree RA affected productivity in regular unpaid activities [0-10 VAS; with 0 indicating no effect and 10 indicating RA completely prevented the subject’s daily activities]). The recall period for questions 2 to 6 is 7 days.

Four main outcomes (expressed in percentages) can be obtained from the WPAI-RA: percentage of work time missed due to RA, percentage of impairment due to RA, percentage of overall work impairment due to RA, percentage of activity impairment due to RA. Note that for subjects who did not work during the 7 days covered by the WPAI-RA, the percent overall work impairment will be equal to the percent of work time missed due to RA.

6.7.1.10. Healthcare Resource Utilization

The Healthcare Resource Utilization Questionnaire is designed to assess healthcare usage during the previous three months across a number of direct medical cost domains.

This questionnaire should be completed by the patient prior to any procedures being performed at the visit, if possible where local language questionnaires are available.

6.7.1.11. Hands, Wrists, and Feet Radiographs

Hands, wrists, and feet radiographs will be taken at the time points indicated in the study procedures table (Appendix 2). Note that if a subject discontinues the study early and has had X-rays performed in the <12 week period prior to discontinuation, these do not need to be repeated at ET. Radiographs performed after enrollment may be done +/-7 days of the scheduled visit.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

All subjects who discontinue from investigational therapy but continue study visits and procedures, should continue to have radiographs performed as outlined in the study procedure table (Appendix 2).

The radiographs will be evaluated through central review by independent blinded assessors using the mTSS.

6.7.1.12 Exploratory patient reported outcomes

Subjects will be asked to rate the effect of their RA on their sexual functioning at the time points indicated in the study procedures table (Appendix 2) where local language questionnaires are available.

The response will be recorded on a 0-100 mm VAS, with 0 indicating `rheumatoid arthritis has no effect on my sexual function` and 100 indicating `rheumatoid arthritis completely inhibits my sexual function`.

Other exploratory patient reported outcomes might be incorporated.

6.7.2 Safety and Tolerability

Adverse events (AEs), physical examinations, vital signs, 12-lead electrocardiograms (ECG) and laboratory assessments (standard hematology, serum/plasma chemistry, and urinalysis) will be collected.

6.7.3 Treatment Satisfaction Questionnaire for Medication (TSQM)

The TSQM will assess the overall level of satisfaction or dissatisfaction with medication subjects are taking. Subjects will be asked to rate their treatment satisfaction at the time points indicated in the study procedures table (Appendix 2). The TSQM will implemented when and where local language TSQM are available.

6.7.4 Clinical Laboratory Evaluations

The hematology and serum chemistry laboratory analyses will be performed at a central laboratory. Reference ranges will be supplied by the central laboratory and will be used by the investigator to assess the laboratory data for clinical significance and pathological changes.

Blood samples will be collected by venipuncture (or optional indwelling catheter for pharmacokinetic sampling days) in the arm at the time points indicated in the study procedures table (Appendix 2). In addition, urine samples for the clinical laboratory assessments will be collected. Subjects only need to be fasted at days were lipid profiling is scheduled.

- Please refer to Appendix 6 for table of Clinical laboratory tests.
The laboratory values outside the normal range will be flagged and clinical relevance will be assessed by the investigator. More frequent sampling as well as additional tests may be performed as deemed necessary by the investigator as indicated.

Note that in the case where clinically significant laboratory test results are a potential reason for discontinuation from the study drug and withdrawal from the study, retesting of the affected parameter(s) should be prompt (within 3 to 7 days) after the investigator has consulted with the medical monitor. A decision regarding subject discontinuation should be made after the results from the retest are available (see Section 3.5 for additional information).

The details of sample handling and shipment instructions will be provided in a separate laboratory manual.

6.7.5. Vital Signs

Vital signs will be measured at the time points indicated in the study procedures table (Appendix 2).

Vital signs should be taken after the subject has been resting for 5 min and will include heart rate, respiratory rate, SBP, DBP, and body temperature.

6.7.6. Physical Examination

A physical examination should be performed at the time points indicated in the study procedures table (Appendix 2).

Any changes from Baseline will be recorded. Height should be measured at Screening only.

At Screening, Week 24, and Week 52 (or at ET), a complete physical examination should be performed. Symptom-directed physical examinations should be performed at all other visits. Weight is measured at all visits.

6.7.7. Other Safety Assessments

6.7.7.1. 12-lead Electrocardiogram

A resting 12-lead ECG should be performed at the time points indicated in the study procedures table (Appendix 2).

The ECG should be obtained after the subject has been resting in the supine position for 5 min and will include heart rate (HR), inter-beat (RR), QRS, uncorrected QT, morphology, and rhythm analysis. QT interval corrected for HR according to Fridericia (QTcF) will be derived during the statistical analysis. ECGs will be interpreted by the investigator for clinical significance and results will be entered into the eCRF.
6.8. **Pharmacokinetics Assessments**

For all subjects, blood samples for PK analysis should be collected at least 30 minutes post study drug dosing at Week 4, prior to study drug dosing at Week 12 and Week 24, and any time at Week 52 or ET.

For subjects who consent to participate in the optional PK substudy (approximately 60 subjects), the daily dose of study drug should be administered under supervision in the clinic (between Week 2 and Week 8, inclusive), and additional PK samples should be collected predose and at 0.5, 1, 2, 3, 4, and 6 hours post dose.

Plasma concentrations of filgotinib and GS-829845 will be analyzed.

Additional analyses (eg, for MTX and its metabolites, adalimumab) may be performed.

6.9. **Biomarker Assessments**

Blood and urine samples will be collected at Day 1, Weeks 4, 8, 12, 24, 36, and week 52 or ET for assessment of markers including but not limited to inflammation, immune status, joint damage, and the JAK-STAT pathway. Genetic predisposition to disease and filgotinib treatment will be assessed in subjects consenting to optional genomic substudy (Section 3.8). Specific information regarding the collection and processing of biomarker samples (if applicable) will be provide to each site in a separate laboratory manual.
7. **ADVERSE EVENTS AND TOXICITY MANAGEMENT**

7.1. **Definitions of Adverse Events, Adverse Reactions, and Serious Adverse Events**

7.1.1. **Adverse Events**

An adverse event (AE) is any untoward medical occurrence in a clinical study subject administered a medicinal product, which does not necessarily have a causal relationship with the treatment. An AE can therefore be any unfavorable and/or unintended sign, symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product. AEs may also include pre- or post-treatment complications that occur as a result of protocol specified procedures, lack of efficacy, overdose, drug abuse/misuse reports, or occupational exposure. Preexisting events that increase in severity or change in nature during or as a consequence of participation in the clinical study will also be considered AEs.

An AE does not include the following:

- Medical or surgical procedures such as surgery, endoscopy, tooth extraction, and transfusion. The condition that led to the procedure may be an adverse event and must be reported.
- Pre-existing diseases, conditions, or laboratory abnormalities present or detected before the screening visit that do not worsen
- Situations where an untoward medical occurrence has not occurred (e.g., hospitalization for elective surgery, social and/or convenience admissions)
- Overdose without clinical sequelae (see Section 7.7.1)
- Any medical condition or clinically significant laboratory abnormality with an onset date before the consent form is signed and not related to a protocol-associated procedure is not an AE. It is considered to be pre-existing and should be documented on the medical history CRF.

7.1.2. **Serious Adverse Events**

A **serious adverse event** (SAE) is defined as an event that, at any dose, results in the following:

- Death
- Life-threatening (Note: The term `life-threatening` in the definition of `serious` refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event that hypothetically might have caused death if it were more severe.)
- In-patient hospitalization or prolongation of existing hospitalization
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

- Persistent or significant disability/incapacity
- A congenital anomaly/birth defect
- A medically important event or reaction: such events may not be immediately life-threatening or result in death or hospitalization but may jeopardize the subject or may require intervention to prevent one of the other outcomes constituting SAEs. Medical and scientific judgment must be exercised to determine whether such an event is a reportable under expedited reporting rules. Examples of medically important events include intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization; and development of drug dependency or drug abuse. For the avoidance of doubt, infections resulting from contaminated medicinal product will be considered a medically important event and subject to expedited reporting requirements.

7.1.3. Clinical Laboratory Abnormalities and Other Abnormal Assessments as Adverse Events or Serious Adverse Events

Laboratory abnormalities without clinical significance are not recorded as AEs or SAEs. However, laboratory abnormalities (eg, clinical chemistry, hematology, and urinalysis) that require medical or surgical intervention or lead to study drug interruption, modification, or discontinuation must be recorded as an AE, as well as an SAE, if applicable. In addition, laboratory or other abnormal assessments (eg, electrocardiogram, x-rays, vital signs) that are associated with signs and/or symptoms must be recorded as an AE or SAE if they meet the definition of an AE or SAE as described in Sections 7.1.1 and 7.1.2 if the laboratory abnormality is part of a syndrome, record the syndrome or diagnosis (eg, anemia), not the laboratory result (ie, decreased hemoglobin).

For specific information on handling of clinical laboratory abnormalities in this study, please refer to Section 7.5.

7.2. Assessment of Adverse Events and Serious Adverse Events

The investigator or qualified subinvestigator is responsible for assessing AEs and SAEs for causality and severity, and for final review and confirmation of accuracy of event information and assessments.

7.2.1. Assessment of Causality for Study Drugs and Procedures

The investigator or qualified subinvestigator is responsible for assessing the relationship to study drug(s) using clinical judgment and the following considerations:

- **No**: Evidence exists that the adverse event has an etiology other than the study drug. For SAEs, an alternative causality must be provided (eg, pre-existing condition, underlying disease, intercurrent illness, or concomitant medication).
- **Yes**: There is reasonable possibility that the event may have been caused by the investigational medicinal product.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

It should be emphasized that ineffective treatment should not be considered as causally related in the context of adverse event reporting.

The relationship to study procedures (e.g., invasive procedures such as venipuncture or biopsy) should be assessed using the following considerations:

- **No**: Evidence exists that the adverse event has an etiology other than the study procedure.
- **Yes**: The adverse event occurred as a result of protocol procedures, (e.g., venipuncture)

### 7.2.2. Assessment of Severity

The severity of AEs will be graded using the modified CTCAE, version 4.03. For each episode, the highest grade attained should be reported.

If a CTCAE criterion does not exist, the investigator should use the grade or adjectives:

- Grade 1 (mild),
- Grade 2 (moderate),
- Grade 3 (severe),
- Grade 4 (life-threatening) or
- Grade 5 (fatal) to describe the maximum intensity of the adverse event. For purposes of consistency with the CTCAE, these intensity grades are defined in Table 7-1 and Appendix 4.

### Table 7-1. Grading of Adverse Event Severity

<table>
<thead>
<tr>
<th>Grade</th>
<th>Adjective</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1</td>
<td>Mild</td>
<td>Asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated</td>
</tr>
<tr>
<td>Grade 2</td>
<td>Moderate</td>
<td>Local or noninvasive intervention indicated; limiting age-appropriate instrumental ADL</td>
</tr>
<tr>
<td>Grade 3</td>
<td>Severe</td>
<td>Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL</td>
</tr>
<tr>
<td>Grade 4</td>
<td>Life-threatening</td>
<td>Urgent intervention indicated</td>
</tr>
<tr>
<td>Grade 5</td>
<td>Death</td>
<td>Death related AE</td>
</tr>
</tbody>
</table>

* Activities of Daily Living (ADL) Instrumental ADL refer to opening preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.
** Self-care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden.

### 7.3. Investigator Requirements and Instructions for Reporting Adverse Events and Serious Adverse Events to Gilead

**Requirements for collection prior to study drug initiation:**

After informed consent, but prior to initiation of study medication, the following types of events should be reported on the case report form (CRF/eCRF): all SAEs and adverse events related to protocol-mandated procedures.
Adverse Events

Following initiation of study medication, collect all AEs, regardless of cause or relationship, until 30-days after last administration of study drug(s) must be reported to the CRF/eCRF database as instructed.

All AEs should be followed up until resolution or until the adverse event is stable, if possible. Gilead Sciences may request that certain AEs be followed beyond the protocol defined follow up period.

Serious Adverse Events

All SAEs, regardless of cause or relationship, that occurs after the subject first consents to participate in the study (ie, signing the informed consent) and throughout the duration of the study, including the protocol-required post treatment follow-up period, must be reported to the CRF/eCRF database and Gilead Drug Safety and Public Health (DSPH) as instructed. This also includes any SAEs resulting from protocol-associated procedures performed after informed consent is signed.

Any SAEs and deaths that occur after the post treatment follow-up visit but within 30-days of the last dose of study drug(s), regardless of causality, should also be reported. Investigators are not obligated to actively seek SAEs after the protocol defined follow up period; however, if the investigator learns of any SAEs that occur after study participation has concluded and the event is deemed relevant to the use of study drug(s), he/she should promptly document and report the event to Gilead DSPH.

- All AEs and SAEs will be recorded in the CRF/eCRF database within the timelines outlined in the CRF/eCRF completion guideline.

Electronic Serious Adverse Event (eSAE) Reporting Process

- Site personnel record all SAE data in the eCRF database and from there transmit the SAE information to Gilead DSPH within 24 hours of the investigator’s knowledge of the event. Detailed instructions can be found in the eCRF completion guidelines.

- If for any reason it is not possible to record the SAE information electronically, ie, the eCRF database is not functioning, record the SAE on the paper serious adverse event reporting form and submit within 24 hours as described above.

  Gilead DSPH: Fax: 1-650-522-5477
  E-mail: Safety_FC@gilead.com

- As soon as it is possible to do so, any SAE reported via paper must be transcribed into the eCRF Database according to instructions in the eCRF completion guidelines.
• If an SAE has been reported via a paper form because the eCRF database has been locked, no further action is necessary.

• All AEs and SAEs will be recorded in the eCRF database within the timelines outlined in the eCRF completion guideline.

• For fatal or life-threatening events, copies of hospital case reports, autopsy reports, and other documents are also to be submitted by e-mail or fax when requested and applicable. Transmission of such documents should occur without personal subject identification, maintaining the traceability of a document to the subject identifiers.

• Additional information may be requested to ensure the timely completion of accurate safety reports.

• Any medications necessary for treatment of the SAE must be recorded onto the concomitant medication section of the subject’s CRF/eCRF and the event description section of the SAE form.

7.4. **Gilead Reporting Requirements**

Depending on relevant local legislation or regulations, including the applicable US FDA Code of Federal Regulations, the EU Clinical Trials Directive (2001/20/EC) and relevant updates, and other country-specific legislation or regulations, Gilead may be required to expedite to worldwide regulatory agencies reports of SAEs, serious adverse drug reactions (SADRs), or suspected unexpected serious adverse reactions (SUSARs). In accordance with the EU Clinical Trials Directive (2001/20/EC), Gilead or a specified designee will notify worldwide regulatory agencies and the relevant IEC in concerned Member States of applicable SUSARs as outlined in current regulations.

Assessment of expectedness for SAEs will be determined by Gilead using reference safety information specified in the investigator’s brochure or relevant local label as applicable.

All investigators will receive a safety letter notifying them of relevant SUSAR reports associated with any study drug(s). The investigator should notify the IRB or IEC of SUSAR reports as soon as is practical, where this is required by local regulatory agencies, and in accordance with the local institutional policy.

7.5. **Clinical Laboratory Abnormalities and Other Abnormal Assessments as Adverse Events or Serious Adverse Events**

Laboratory abnormalities are usually not recorded as AEs or SAEs. However, laboratory abnormalities (eg, clinical chemistry, hematology, and urinalysis) independent of the underlying medical condition that require medical or surgical intervention or lead to investigational medicinal product interruption or discontinuation must be recorded as an AE, as well as an SAE, if applicable. In addition, laboratory or other abnormal assessments (eg, electrocardiogram, X-rays, vital signs) that are associated with signs and/or symptoms must
be recorded as an AE or SAE if they meet the definition of an AE (or SAE) as described in Sections 7.1.1 and 7.1.2. If the laboratory abnormality is part of a syndrome, record the syndrome or diagnosis (i.e., anemia) not the laboratory result (i.e., decreased hemoglobin).

Adverse events will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA). Severity should be recorded and graded according to the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03, which can be found at:


For AEs associated with laboratory abnormalities, the event should be graded on the basis of the clinical severity in the context of the underlying conditions; this may or may not be in agreement with the grading of the laboratory abnormality.

All clinical and clinically significant laboratory toxicities will be managed according to uniform guidelines detailed in Appendix 3 and as outlined below.

7.5.1. Grades 1 and 2 Laboratory Abnormality or Clinical Event

Continue study drug at the discretion of the investigator.

7.5.2. Grades 3 Laboratory Abnormality or Clinical Event

- For Grade 3 clinically significant laboratory abnormality or clinical event, investigational medicinal product may be continued if the event is considered to be unrelated to investigational medicinal product.

- For a Grade 3 clinical event, or clinically significant laboratory abnormality confirmed by repeat testing, that is considered to be related to investigational medicinal product, investigational medicinal product should be withheld until the toxicity returns to Grade 2.

- If a laboratory abnormality recurs to Grade 3 following re-challenge with investigational medicinal product and is considered related to investigational medicinal product, then investigational medicinal product should be permanently discontinued and the subject managed according to local clinical practice. Recurrence of laboratory abnormalities considered unrelated to investigational medicinal product may not require permanent discontinuation.

7.5.3. Grades 4 Laboratory Abnormality or Clinical Event

- For a Grade 4 clinical event or clinically significant Grade 4 laboratory abnormality confirmed by repeat testing that is considered related to investigational medicinal product, investigational medicinal product should be permanently discontinued and the subject managed according to local clinical practice. The subject should be followed as clinically indicated until the laboratory abnormality returns to baseline or is otherwise explained,
whichever occurs first. A clinically significant Grade 4 laboratory abnormality that is not confirmed by repeat testing should be managed according to the algorithm for the new toxicity grade.

Investigational medicinal product may be continued without dose interruption for a clinically non-significant Grade 4 laboratory abnormality (e.g., Grade 4 CK after strenuous exercise or triglyceride elevation that is nonfasting or that can be medically managed) or a clinical event considered unrelated to investigational medicinal product.

7.6. Toxicity Management

Treatment-emergent toxicities will be noted by the investigator and brought to the attention of the Gilead Sciences medical monitor, who will have a discussion with the investigator and decide the appropriate course of action. Whether or not considered treatment-related, all subjects experiencing AEs must be monitored periodically until symptoms subside, any abnormal laboratory values have resolved or returned to baseline levels or they are considered irreversible, or until there is a satisfactory explanation for the changes observed.

Grade 3 or 4 clinically significant laboratory abnormalities should be managed as outlined in Appendix 3 and Sections 7.5.2 and 7.5.3.

Any questions regarding toxicity management should be directed to the Gilead Sciences Medical Monitor.

7.7. Special Situations Reports

7.7.1. Definitions of Special Situations

Special situation reports include all reports of medication error, abuse, misuse, overdose, reports of AEs associated with product complaints, occupational exposure with an AE, pregnancy reports (regardless of an associated AE), and AEs in an infant following potential exposure from breastfeeding.

Medication error is any unintentional error in the prescribing, dispensing, or administration of a medicinal product while in the control of the health care provider, subject, or consumer.

Abuse is defined as persistent or sporadic intentional excessive use of a medicinal product by a subject.

Misuse is defined as any intentional and inappropriate use of a medicinal product that is not in accordance with the protocol instructions or the local prescribing information.

An overdose is defined as an accidental or intentional administration of a quantity of a medicinal product given per administration or cumulatively which is above the maximum recommended dose as per protocol or in the product labelling (as it applies to the daily dose of the subject in question). In cases of a discrepancy in drug accountability, overdose will be established only when it is clear that the subject has taken the excess dose(s). Overdose cannot be established when the subject cannot account for the discrepancy except in cases in which the investigator has reason to suspect that the subject has taken the additional dose(s).
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc. Final Amendment 1

Product complaint is defined as complaints arising from potential deviations in the manufacture, packaging, or distribution of the medicinal product.

Occupational exposure is defined as exposure to a medicinal product as a result of one's professional or non-professional occupation.

7.7.2. Instructions for Reporting Special Situations

7.7.2.1. Instructions for Reporting Pregnancies

The investigator should report pregnancies in female study subjects that are identified after initiation of study drug and throughout the study, including the post study drug follow-up period, to Gilead DSPH using the pregnancy report form within 24 hours of becoming aware of the pregnancy.

Refer to Section 7.3 and the CRF/eCRF completion guidelines for full instructions on the mechanism of pregnancy reporting.

The pregnancy itself is not considered an AE nor is an elective abortion to terminate a pregnancy without medical reasons.

Any premature termination of pregnancy (e.g., a spontaneous abortion, an induced therapeutic abortion due to complications or other medical reasons) must be reported within 24 hours as an SAE. The underlying medical reason for this procedure should be recorded as the AE term.

A spontaneous abortion is always considered to be an SAE and will be reported as described in Sections 7.1.1 and 7.1.2. Furthermore, any SAE occurring as an adverse pregnancy outcome post study must be reported to Gilead DSPH.

The subject should receive appropriate monitoring and care until the conclusion of the pregnancy. The outcome should be reported to Gilead using the pregnancy outcome report form. If the end of the pregnancy occurs after the study has been completed, the outcome should be reported directly to Gilead DSPH. Gilead DSPH contact information is as follows: Email: Safety_FC@gilead.com and Fax: +1 (650) 522-5477.

Pregnancies of female partners of male study subjects exposed to study drugs must also be reported and relevant information should be submitted to Gilead DSPH using the pregnancy and pregnancy outcome forms within 24 hours. Monitoring of the subject should continue until the conclusion of the pregnancy. If the end of the pregnancy occurs after the study has been completed, the outcome should be reported directly to Gilead DSPH, fax number +1 650 522-5477 or email Safety_FC@gilead.com.

Refer to Appendix 5 for Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Requirements.
7.7.2.2. Reporting Other Special Situations

All other special situation reports must be reported on the special situations report form and forwarded to Gilead DSPh within 24 hours of the investigator becoming aware of the situation. These reports must consist of situations that involve study drug(s) and/or Gilead concomitant medications, but do not apply to non-Gilead concomitant medications.

Special situations involving non-Gilead concomitant medications does not need to be reported on the special situations report form; however, for special situations that result in AEs due to a non-Gilead concomitant medication, the AE should be reported on the AE form.

Any inappropriate use of concomitant medications prohibited by this protocol should not be reported as ‘misuse,’ but may be more appropriately documented as a protocol deviation.

Refer to Section 7.3 and the CRF/eCRF completion guidelines for full instructions on the mechanism of special situations reporting.

All clinical sequelae in relation to these special situation reports will be reported as AEs or SAEs at the same time using the AE CRF/eCRF and/or the SAE report form. Details of the symptoms and signs, clinical management, and outcome will be reported, when available.
8. STATISTICAL CONSIDERATIONS

8.1. Analysis Objectives and Endpoints

8.1.1. Analysis Objectives

The primary objective of this study is as follows:

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving ACR20 at Week 12

The secondary objectives of this study are as follows:

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving Disease Activity Score for 28 joint count using CRP (DAS28 [CRP]) ≤3.2 at Week 12
- To evaluate the effect of filgotinib versus placebo on physical function as measured by change from Baseline in the HAQ-DI score at Week 12
- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP) <2.6 at Week 24
- To evaluate the effects of filgotinib versus placebo on preservation of joint structure as measured by change from Baseline in the van der Heijde mTSS at Week 24
- To evaluate the effects of filgotinib versus adalimumab for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28 (CRP) ≤3.2 at Week 12
- To evaluate the safety and tolerability of filgotinib
- To evaluate the effects of filgotinib on work productivity, fatigue, and general quality of life as measured by SF-36, FACIT-Fatigue, EQ-5D and WPAI-RA

The exploratory objectives of this study include:

- To characterize the PK of filgotinib and its metabolite (GS-829845)
- To characterize the association of host genetics and other markers with disease severity, disease progression and treatment response to filgotinib in subjects with rheumatoid arthritis
- To evaluate the effects of filgotinib on healthcare resource utilization and other patient reported outcomes
8.1.2. **Primary Endpoint**

The primary endpoint is the proportion of subjects who achieve an ACR20 response at Week 12.

8.1.3. **Secondary Endpoints**

The key secondary endpoints are:

- The proportion of subjects who achieve DAS28 (CRP) ≤ 2 at Week 12
- Change from Baseline in the HAQ-DI score at Week 12
- The proportion of subjects who achieve DAS28 (CRP) < 2.6 at Week 24
- Change from Baseline in mTSS at Week 24

Other secondary endpoints include:

- The proportion of subjects who achieve ACR50 and ACR70 at Weeks 4, 12, 24 and 52, ACR20 at Weeks 4, 24, and 52, and ACR20/50/70 over time from Day 1 through Week 52
- Change from Baseline in individual components of the ACR response at Weeks 4, 12, 24, and 52 and over time from Day 1 through Week 52
- The proportion of subjects who achieve change in HAQ-DI of ≤ 0.22 at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in DAS28 (CRP) at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28 (CRP) ≤ 2 at Weeks 4, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28 (CRP) < 2.6 at Weeks 4, 12, and 52, and over time from Day 1 through Week 52
- ACR-N and EULAR response at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in CDAI at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in SDAI at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 24
- Change from Baseline in the mTSS at Week 52
- The proportion of subjects with no radiographic progression from Baseline at Week 24 and 52

- Absolute value and change from Baseline in SF-36, FACIT-Fatigue, and the EQ-5D over time at Weeks 4, 12, 24 and 52, and over time from Day 1 through Week 52

- Absolute value and change from Baseline in WPAI-RA at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52

8.2. Analysis Conventions

8.2.1. Analysis Sets

8.2.1.1. All Randomized

The all randomized analysis set includes all subjects who are randomized in the study. This is the primary analysis set for by-subject listings.

8.2.1.2. Efficacy

8.2.1.2.1. Full Analysis Set (FAS)

The primary analysis set for efficacy analyses will be the Full Analysis Set (FAS), which includes all randomized subjects who received at least one dose of study drug.

8.2.1.2.2. Per-Protocol (PP) Analysis Set

The secondary analysis set for efficacy analyses will be the Per-Protocol (PP) Analysis set, which includes all subjects in the Full Analysis Set who have not committed any major protocol violation, including the violation of key entry criteria.

8.2.1.3. Safety

The primary analysis set for safety analyses will be the Safety Analysis Set, which includes all subjects who received at least one dose of study drug.

8.2.1.4. Pharmacokinetic

8.2.1.4.1. Pharmacokinetic (PK) Substudy Analysis Set

The primary analysis set for intensive PK analyses will be the PK substudy analysis set, which includes all subjects in the Safety Analysis Set who have enrolled into the PK Substudy, and have intensive concentration data to provide interpretable results for the specific parameters of interest for the analyte under evaluation.
8.2.1.4.2. Pharmacokinetic (PK) Analysis Set

The primary analysis set for general PK analyses will be the PK analysis set, which includes all subjects in the Safety Analysis Set who have at least 1 nonmissing concentration data for filgotinib and/or its metabolite GS-829845.

8.3. Data Handling Conventions

PK concentration values and PK parameter values below the limit of quantitation (BLQ) will be presented as "BLQ" in the data listings. BLQ values that occur prior to the first dose will be treated as 0, BLQ values at all other time points will be treated as 1/2 of the lower limit of quantitation (LLOQ).

Laboratory data that are continuous in nature but are less than the lower limit of quantitation or above the upper limit of quantitation will be imputed to the value of the lower or upper limit minus or plus one significant digit, respectively (eg, if the result of a continuous laboratory test is < 20, a value of 19 will be assigned; if the result of a continuous laboratory test is < 20.0, a value of 19.9 will be assigned).

8.4. Demographic Data and Baseline Characteristics

Demographic and baseline characteristics will be summarized using standard descriptive statistics including sample size, mean, SD, median, Q1, Q3, minimum, and maximum for continuous variables and numbers and percentages of subjects for categorical variables.

Demographic data will include sex, race, ethnicity, and age.

Baseline characteristics may include prior exposure to bDMARDs, RF status, anti-CCP Ab status, DAS28(CRP), HAQ-DI, mTSS, SDAI, CDAI, and other variables of interest.

8.5. Efficacy Analysis

8.5.1. Primary Analysis

The primary endpoint for the study is the proportion of subjects who achieve an ACR20 response at Week 12. The primary analysis will consist of a superiority test of filgotinib 200 mg compared to placebo based on the ACR20 response rate at Week 12. Cochran-Mantel-Haenszel (CMH) approach adjusting for the randomization stratification factors will be used for the hypothesis testing at the 2-sided 0.05-level. Subjects who do not have sufficient measurements to establish efficacy at Week 12 will be considered as failures (ie, non-responder imputation [NRI]). Sensitivity analyses will be conducted and described in the statistical analysis plan (SAP).
8.5.2. Secondary Analyses

The following hypothesis testing will commence after the primary analysis reaches statistical
significance, and will be tested according to the hierarchical testing principle at the 2-sided
0.05-level. If a null hypothesis is not rejected, formal sequential testing will be stopped and only
nominal significance will be reported for the remaining hypotheses:

1) Superiority of filgotinib 200 mg compared to placebo based on the response rate of DAS28
   (CRP) ≤ 2.6 at Week 12.

2) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in
   HAQ-DI at Week 12.

3) Superiority of filgotinib 200 mg compared to placebo based on the response rate of DAS28
   (CRP) < 2.6 at Week 24.

4) Superiority of filgotinib 100 mg compared to placebo based on ACR20 response rate at
   Week 12.

5) Superiority of filgotinib 100 mg compared to placebo based on the response rate of DAS28
   (CRP) ≤ 2.6 at Week 12.

6) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in
   HAQ-DI at Week 12.

7) Superiority of filgotinib 100 mg compared to placebo based on the response rate of DAS28
   (CRP) < 2.6 at Week 24.

8) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in
   mTSS at Week 24.

9) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in
   mTSS at Week 24.

10) Non-inferiority of filgotinib 200 mg compared to adalimumab based on the response rate of
    DAS28 (CRP) ≤ 2.6 at Week 12.

11) Non-inferiority of filgotinib 100 mg compared to adalimumab based on the response rate
    DAS28 (CRP) ≤ 2.6 at Week 12.

For categorical endpoints (response rate of DAS28 [CRP] ≤ 2.6 and DAS28 [CRP] < 2.6), the same
CMH approach with NRI as in the primary analysis will be adopted. For continuous endpoints
(HAQ-DI and mTSS), mixed-effects model for repeated measures (MMRM) will be used to
evaluate treatment effect on change score from Baseline, with treatment, visit, stratification
factors and baseline value included as fixed effects and subject being the random effect. Missing
change scores in HAQ-DI and mTSS due to early withdrawal or treatment reassignment will not
be otherwise imputed using the MMRM approach. Sensitivity analyses will be conducted by
imputing missing data via, for example, last observation carried forward (LOCF) and multiple
imputation (MI) methods.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

For the non-inferiority test, the approach proposed by Liu et al. 2014 will be used to demonstrate that each filgotinib dose preserves more than 50% of the effect of adalimumab on the response rate of DAS28 (CRP)Hb.2 at Week 12. Let \( \pi_T, \pi_C, \) and \( \pi_P \) denote the true response rates of DAS28 (CRP)Hb.2 for filgotinib, adalimumab, and placebo, respectively, at Week 12, with corresponding variances \( \sigma_T^2, \sigma_C^2, \) and \( \sigma_P^2 \). The non-inferiority null hypothesis is

\[
H_{0,NI}: \frac{\pi_T - \pi_P}{\pi_C - \pi_P} \leq 0.5.
\]

According to Liu et al. 2014, non-inferiority (filgotinib preserves more than 50% of the effect of adalimumab) at 2-sided 0.05 level will be claimed if \( Z_{NI} > z_0 \), where

\[
Z_{NI} = \frac{\hat{\pi}_T - 0.5\hat{\pi}_C - (1-0.5)\hat{\pi}_P}{\sqrt{\hat{\sigma}_T^2 + 0.5^2\hat{\sigma}_C^2 + (1-0.5)^2\hat{\sigma}_P^2}}
\]

and \( z_0 \) is the 97.5 quantile of the standard normal distribution.

Note that the `hat` (\( \hat{\cdot} \)) denotes the estimated values of each parameter at Week 12.

If non-inferiority of at least one filgotinib dose compared to adalimumab based on the response rate of DAS28 (CRP)Hb.2 at Week 12 is demonstrated, then a superiority test of the respective filgotinib dose compared to adalimumab will be conducted.

For other secondary endpoints listed under Section 8.1.3, summary statistics will be provided by treatment group. Differences across treatment groups will be summarized and treatment comparisons may be performed. More details on efficacy analyses will be described in the SAP.

8.6. Safety Analysis

All safety analyses will be performed using the safety analysis set.

Safety will be evaluated by assessment of clinical laboratory tests, physical examinations, vital signs measurements at various time points during the study, and by the documentation of AEs.

All safety data collected on or after the first dose of study drug administration up to 30 days after the last dose of study drug, unless specified otherwise, will be summarized by treatment group according to the study drug received.

8.6.1. Extent of Exposure

A subject’s extent of exposure to study drug will be generated from the study drug administration page of the eCRF. Exposure data will be summarized by treatment group.

Duration of exposure to study drug will be expressed as the number of weeks between the first and last dose of the study drug, inclusive, regardless of temporary interruptions in study drug administration and summarized by treatment group.
8.6.2. **Adverse Events**

Clinical and laboratory adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). System Organ Class (SOC), High-Level Group Term (HLGT), High-Level Term (HLT), Preferred Term (PT), and Lower-Level Term (LLT) will be attached to the clinical database.

Treatment-Emergent Adverse Events (TEAEs) are:

- Any AEs with an onset date of on or after the study drug start date and no later than 30 days after permanent discontinuation of study drug or
- Any AEs leading to premature discontinuation of study drug.

Summaries (number and percentage of subjects) of TEAEs by SOC and PT will be provided by treatment group. TEAEs will also be summarized by relationship to study drug and severity. In addition, TEAEs leading to premature discontinuation of study drug will be summarized and listed.

8.6.3. **Laboratory Evaluations**

Selected laboratory data will be summarized (n, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment group and study visit along with corresponding change from Baseline. The incidence of treatment-emergent graded laboratory abnormalities will be summarized similarly.

Graded laboratory abnormalities will be defined using CTCAE 4.03 grading scale Appendix 4.

8.7. **Pharmacokinetic Analysis**

Plasma concentrations of filgotinib and its metabolite (GS-829845) will be listed and summarized for all subjects by treatment using descriptive statistics (eg, sample size, arithmetic mean, geometric mean, % coefficient of variation, standard deviation, median, minimum, and maximum). For subjects enrolled in the optional PK substudy, pharmacokinetic parameters of filgotinib and its metabolite GS-829845 will be listed and summarized by treatment; and plasma concentrations of the filgotinib and GS-829845 over time will be plotted in semi logarithmic and linear formats as mean ± standard deviation and median (Q1, Q3) by treatment.

Exposure-response analyses may be performed.

8.8. **Biomarker Analysis**

Exploratory analyses may be performed to evaluate the association of each biomarker or combination of biomarkers with clinical outcomes, the modulation of biomarkers related to mechanism of action and disease progression, and biomarker or combination of biomarkers predictive of treatment response.
Exploratory biomarkers analyses that may enhance the understanding of the biological effects, the mechanism of action, or safety, may be performed. Biomarker objectives may be further described and updated based on evolving scientific knowledge of filgotinib. If an exploratory biomarker analysis is to be performed, biomarker analysis plan, with details on objectives and analysis methods, will be issued prior to the actual data analysis.

8.9.  Sample Size

Sample size is determined based on the superiority test of filgotinib 200 mg compared to placebo based on the change from Baseline in mTSS at Week 24. When assuming a difference of 0.4 between filgotinib and placebo on change from Baseline in mTSS at Week 24 and a common standard deviation of 1.85, 450 subjects in each of the filgotinib 200 mg group and placebo control group are required to obtain 90% power at a 2-sided 0.05-level.

A sample size of 450 subjects in each of the filgotinib groups and placebo control group will provide over 95% power to detect an increase in ACR20 response rate of 45% to 65% between the placebo control group and the filgotinib group respectively, using a 2-sided 0.05-level test.

Based on \cite{Liu et al 2014}, 450 subjects in each of the filgotinib 200 mg group and placebo group, and 300 subjects in the adalimumab group, will provide over 90% power at 2-sided 0.05 significance level to demonstrate that filgotinib 200 mg preserves more than 50% of the effect of adalimumab with respect to the response rate of DAS28 (CRP)\textsubscript{H}.2 at Week 12, assuming both filgotinib 200 mg and adalimumab groups have similar response rates of DAS28 (CRP)\textsubscript{H}.2. Given this study has a placebo group, assay sensitivity can be established through a direct comparison of adalimumab to placebo.

In summary, the total sample size will be 1650 (450 for the filgotinib 200 mg, filgotinib 100 mg, placebo groups, and 300 for the active comparator group).

8.10.  Data Monitoring Committee

An external multidisciplinary DMC will review the progress of the study and perform interim reviews of safety data and provide recommendation to Gilead whether the nature, frequency, and severity of adverse effects associated with study treatment warrant the early termination of the study in the best interests of the participants, whether the study should continue as planned, or the study should continue with modifications.

The initial reviews will be conducted after approximately 100 subjects enrolled complete through 12 weeks of treatment. The DMC’s specific activities will be defined by a mutually agreed charter, which will define the DMC’s membership, conduct and meeting schedule.

While the DMC will be asked to advise Gilead regarding future conduct of the study, including possible early study termination, Gilead retains final decision-making authority on all aspects of the study.
8.11. **Cardiovascular Endpoint Adjudication Committee**

A Cardiovascular Endpoint Adjudication Committee (CV-EAC) consisting of at least 2 cardiologists and governed by a Charter will be set up to perform adjudication of Major Adverse Cardiovascular Events occurring during the study. The adjudication of these events will be performed in a blinded fashion for the purposes of data analysis, and not for monitoring of subject safety.

8.12. **Analysis Schedule**

The primary and secondary analyses will be conducted after all subjects either complete their Week 24 visit or prematurely discontinue from the study. The final analysis will be performed when all subjects complete the study or prematurely discontinue from the study.
9. RESPONSIBILITIES

9.1. Investigator Responsibilities

9.1.1. Good Clinical Practice

The investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki (as amended in Edinburgh, Tokyo, Venice, Hong Kong, and South Africa), International Conference on Harmonisation (ICH) guidelines, or with the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the study subject. These standards are consistent with the European Union Clinical Trials Directive 2001/20/EC and Good Clinical Practice Directive 2005/28/EC.


The investigator and all applicable subinvestigators will comply with 21 CFR, Part 54, 1998, providing documentation of their financial interest or arrangements with Gilead, or proprietary interests in the investigational drug under study. This documentation must be provided prior to the investigator’s (and any subinvestigator’s) participation in the study. The investigator and subinvestigator agree to notify Gilead of any change in reportable interests during the study and for 1 year following completion of the study. Study completion is defined as the date when the last subject completes the protocol-defined activities.

9.1.2. Institutional Review Board (IRB)/Independent Ethics Committee (IEC) Review and Approval

The investigator (or sponsor as appropriate according to local regulations) will submit this protocol, informed consent form, and any accompanying material to be provided to the subject (such as advertisements, subject information sheets, or descriptions of the study used to obtain informed consent) to an IRB/IEC. The investigator will not begin any study subject activities until approval from the IRB/IEC has been documented and provided as a letter to the investigator.

Before implementation, the investigator will submit to and receive documented approval from the IRB/IEC any modifications made to the protocol or any accompanying material to be provided to the subject after initial IRB/IEC approval, with the exception of those necessary to reduce immediate risk to study subjects.

9.1.3. Informed Consent

The investigator is responsible for obtaining written informed consent from each individual participating in this study after adequate explanation of the aims, methods, objectives, and potential hazards of the study and before undertaking any study-related procedures.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

investigator must use the most current IRB/IEC approved consent form for documenting written informed consent. Each informed consent (or assent as applicable) will be appropriately signed and dated by the subject and the person conducting the consent discussion, and also by an impartial witness if required by IRB/IEC or local requirements. The consent form will inform subjects about genomic testing and sample retention, and their right to receive clinically relevant genomic analysis results.

9.1.4. Confidentiality

The investigator must assure that subjects’ anonymity will be strictly maintained and that their identities are protected from unauthorized parties. Only subject initials, date of birth, another unique identifier (as allowed by local law) and an identification code will be recorded on any form or biological sample submitted to the Sponsor, IRB/IEC, or laboratory. Laboratory specimens must be labeled in such a way as to protect subject identity while allowing the results to be recorded to the proper subject. Refer to specific laboratory instructions. NOTE: The investigator must keep a screening log showing codes, names, and addresses for all subjects screened and for all subjects enrolled in the trial. Subject data will be processed in accordance with all applicable regulations.

The investigator agrees that all information received from Gilead, including but not limited to the investigator brochure, this protocol, CRF/eCRF, the study drug, and any other study information, remain the sole and exclusive property of Gilead during the conduct of the study and thereafter. This information is not to be disclosed to any third party (except employees or agents directly involved in the conduct of the study or as required by law) without prior written consent from Gilead. The investigator further agrees to take all reasonable precautions to prevent the disclosure by any employee or agent of the study site to any third party or otherwise into the public domain.

9.1.5. Study Files and Retention of Records

The investigator must maintain adequate and accurate records to enable the conduct of the study to be fully documented and the study data to be subsequently verified. These documents should be classified into at least the following two categories: (1) investigator’s study file, and (2) subject clinical source documents.

The investigator’s study file will contain the protocol/amendments, CRF and query forms, IRB/IEC and governmental approval with correspondence, informed consent, drug records, staff curriculum vitae and authorization forms, and other appropriate documents and correspondence.

The required source data should include sequential notes containing at least the following information for each subject:

- Subject identification (name, date of birth, gender);
- Documentation that subject meets eligibility criteria, ie, history, physical examination, and confirmation of diagnosis (to support inclusion and exclusion criteria);
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

- Documentation of the reason(s) a consented subject is not enrolled
- Participation in study (including study number);
- Study discussed and date of informed consent;
- Dates of all visits;
- Documentation that protocol specific procedures were performed;
- Results of efficacy parameters, as required by the protocol;
- Start and end date (including dose regimen) of study drug(s), including dates of dispensing and return;
- Record of all adverse events and other safety parameters (start and end date, and including causality and severity);
- Concomitant medication (including start and end date, dose if relevant; dose changes);
- Date of study completion and reason for early discontinuation, if it occurs.

All clinical study documents must be retained by the investigator until at least 2 years or according to local laws, whichever is longer, after the last approval of a marketing application in an ICH region (i.e., United States, Europe, or Japan) and until there are no pending or planned marketing applications in an ICH region; or, if no application is filed or if the application is not approved for such indication, until 2 years after the investigation is discontinued and regulatory authorities have been notified. Investigators may be required to retain documents longer if specified by regulatory requirements, by local regulations, or by an agreement with Gilead. The investigator must notify Gilead before destroying any clinical study records.

Should the investigator wish to assign the study records to another party or move them to another location, Gilead must be notified in advance.

If the investigator cannot provide for this archiving requirement at the study site for any or all of the documents, special arrangements must be made between the investigator and Gilead to store these records securely away from the site so that they can be returned sealed to the investigator in case of an inspection. When source documents are required for the continued care of the subject, appropriate copies should be made for storage away from the site.

9.1.6. Case Report Forms

For each subject consented, an eCRF will be completed by an authorized study staff member whose training for this function is documented according to study procedures. eCRF should be completed on the day of the subject visit to enable the sponsor to perform central monitoring of safety data. The Eligibility Criteria eCRF should be completed only after all data related to
eligibility have been received. Subsequent to data entry, a study monitor will perform source
data verification within the EDC system. Original entries as well as any changes to data fields
will be stored in the audit trail of the system. Prior to database lock (or any interim time points
as described in the clinical data management plan), the investigator will use his/her log in
credentials to confirm that the forms have been reviewed, and that the entries accurately reflect
the information in the source documents. The eCRF captures the data required per the protocol
schedule of events and procedures. System-generated or manual queries will be issued to the
investigative site staff as data discrepancies are identified by the monitor or internal Gilead staff,
who routinely review the data for completeness, correctness, and consistency. The site
coordinator is responsible for responding to the queries in a timely manner, within the system,
either by confirming the data as correct or updating the original entry, and providing the reason
for the update (eg data entry error). At the conclusion of the trial, Gilead will provide the site
with a read-only archive copy of the data entered by that site. This archive must be stored in
accordance with the records retention requirements outlined in Section 9.1.5.

9.1.7. Investigational Medicinal Product Accountability and Return

Where possible, study drug(s) should be destroyed at the site. At the start of the study, the study
monitor will evaluate each study center’s study drug disposal procedures and provide appropriate
instruction for disposal or return of unused study drug supplies. If the site has an appropriate
standard operating procedure (SOP) for drug destruction as determined by Gilead Sciences, the
site may destroy used (empty or partially empty) and unused study drug supplies as long as
performed in accordance with the site’s SOP. This can occur only after the study monitor has
performed drug accountability during an on-site monitoring visit.

A copy of the site’s Study Drug Disposal SOP or written procedure (signed and dated by the PI
or designee) will be obtained for Gilead site files. If the site does not have acceptable
procedures in place, arrangements will be made between the site and Gilead Sciences
(or Gilead Sciences' representative) for return of unused study drug supplies.

If study drug is destroyed on site, the investigator must maintain accurate records for all study
drugs destroyed. Upon study completion, copies of the study drug accountability records must be
filed at the site. Another copy will be returned to Gilead.

The study monitor will review study drug supplies and associated records at periodic intervals.

9.1.8. Inspections

The investigator will make available all source documents and other records for this trial to
Gilead’s appointed study monitors, to IRBs/IECs, or to regulatory authority or health authority
inspectors.

9.1.9. Protocol Compliance

The investigator is responsible for ensuring the study is conducted in accordance with the
procedures and evaluations described in this protocol.
9.2. Sponsor Responsibilities

9.2.1. Protocol Modifications

Protocol modifications, except those intended to reduce immediate risk to study subjects, may be made only by Gilead. The investigator must submit all protocol modifications to the IRB/IEC in accordance with local requirements and receive documented IRB/IEC approval before modifications can be implemented.

9.2.2. Study Report and Publications

A clinical study report (CSR) will be prepared and provided to the regulatory agency(ies). Gilead will ensure that the report meets the standards set out in the ICH Guideline for Structure and Content of Clinical Study Reports (ICH E3). Note that an abbreviated report may be prepared in certain cases.

Investigators in this study may communicate, orally present, or publish in scientific journals or other scholarly media only after the following conditions have been met:

the results of the study in their entirety have been publicly disclosed by or with the consent of Gilead in an abstract, manuscript, or presentation form or the study has been completed at all study sites for at least 2 years.

The investigator will submit to Gilead any proposed publication or presentation along with the respective scientific journal or presentation forum at least 30 days before submission of the publication or presentation.

No such communication, presentation, or publication will include Gilead’s confidential information (see Section 9.1.4).

The investigator will comply with Gilead’s request to delete references to its confidential information (other than the study results) in any paper or presentation and agrees to withhold publication or presentation for an additional 60 days in order to obtain patent protection if deemed necessary.

9.3. Joint Investigator/Sponsor Responsibilities

9.3.1. Payment Reporting

Investigators and their study staff may be asked to provide services performed under this protocol, eg attendance at Investigator’s Meetings. If required under the applicable statutory and regulatory requirements, Gilead will capture and disclose to Federal and State agencies any expenses paid or reimbursed for such services, including any clinical trial payments, meal, travel expenses or reimbursements, consulting fees, and any other transfer of value.
9.3.2. **Access to Information for Monitoring**

In accordance with regulations and guidelines, the study monitor must have direct access to the investigator’s source documentation in order to verify the accuracy of the data recorded in the CRF/eCRF.

The monitor is responsible for routine review of the CRF/eCRF at regular intervals throughout the study to verify adherence to the protocol and the completeness, consistency, and accuracy of the data being entered on them. The monitor should have access to any subject records needed to verify the entries on the CRF/eCRF. The investigator agrees to cooperate with the monitor to ensure that any problems detected through any type of monitoring (central, on site) are resolved.

9.3.3. **Access to Information for Auditing or Inspections**

Representatives of regulatory authorities or of Gilead may conduct inspections or audits of the clinical study. If the investigator is notified of an inspection by a regulatory authority the investigator agrees to notify the Gilead medical monitor immediately. The investigator agrees to provide to representatives of a regulatory agency or Gilead access to records, facilities, and personnel for the effective conduct of any inspection or audit.

9.3.4. **Study Discontinuation**

Both the sponsor and the investigator reserve the right to terminate the study at any time. Should this be necessary, both parties will arrange discontinuation procedures and notify the appropriate regulatory authority(ies), IRBs, and IECs. In terminating the study, Gilead and the investigator will assure that adequate consideration is given to the protection of the subjects’ interests.
10. REFERENCES


Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.


11. APPENDICES

Appendix 1. Investigator Signature Page
Appendix 2. Study Procedures Table
Appendix 3. Management of Clinical and Laboratory Adverse Events
Appendix 4. Common Terminology Criteria for Adverse Events (CTCAE) v4.03
Appendix 5. Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Methods
Appendix 6. Laboratory Assessment Table
Appendix 7. List of Joints to be Evaluated (66/68 Joint Count)
Appendix 8. The 2010 American College of Rheumatology “European League Against Rheumatism Collaborative Initiative Classification Criteria for Rheumatoid Arthritis (Aletaha et al 2010)
Appendix 9. American College of Rheumatology Response Evaluations/ Preliminary Definition of Improvement in Rheumatoid Arthritis (Felson et al 1995)
Appendix 10. Disease Activity Score (DAS28) (Prevoo et al 1995)
Appendix 11. Procedures and Specifications
Appendix 1. Investigator Signature Page

GILEAD SCIENCES, INC.

STUDY ACKNOWLEDGEMENT

A Randomized, Double-blind, Placebo- and Active-controlled, Multicenter, Phase 3 Study to Assess the Efficacy and Safety of Filgotinib Administered for 52 weeks in Combination with Methotrexate to Subjects with Moderately to Severely Active Rheumatoid Arthritis Who Have an Inadequate Response to Methotrexate


This protocol has been approved by Gilead Sciences, Inc. The following signature documents this approval.

FRANZISKA HATZKIES, MD
Name (Printed)
Author

06 - July - 2016
Date

INVESTIGATOR STATEMENT

I have read the protocol, including all appendices, and I agree that it contains all necessary details for me and my staff to conduct this study as described. I will conduct this study as outlined herein and will make a reasonable effort to complete the study within the time designated.

I will provide all study personnel under my supervision copies of the protocol and access to all information provided by Gilead Sciences, Inc. I will discuss this material with them to ensure that they are fully informed about the drugs and the study.

Principal Investigator Name (Printed)  Signature

Date  Site Number

CONFIDENTIAL  Page 100  05 July 2016
## Appendix 2. Study Procedures Table

<table>
<thead>
<tr>
<th>EVENT</th>
<th>Screening</th>
<th>Treatment Period</th>
<th>Follow Up&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visits</td>
<td>Day 1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>W2</td>
<td>W4</td>
</tr>
<tr>
<td>Visit window (Days)</td>
<td>(±3 days)</td>
<td>(±3 days)</td>
<td>(±3 days)</td>
</tr>
<tr>
<td>Informed Consent</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion/Exclusion Criteria</td>
<td>X</td>
<td>X&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Demographics and Baseline/ Disease Characteristics</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA diagnosis and Prior RA Treatment, Medical History&lt;sup&gt;f&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete Physical Examination&lt;sup&gt;g&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TB test and Chest X-ray&lt;sup&gt;h&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF/anti-CCP Ab</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Post Treatment Week 4

<sup>b</sup> Day 1 refers to the first day of the treatment period.

<sup>c</sup> W52/ET indicates the 52nd week of the treatment period.

<sup>e</sup> X signifies a required activity.

<sup>f</sup> RA diagnosis and Prior RA Treatment, Medical History includes the diagnosis of rheumatoid arthritis and any prior treatment or medical history.

<sup>g</sup> Complete Physical Examination includes a comprehensive physical examination.

<sup>h</sup> TB test and Chest X-ray refers to tests for tuberculosis and chest X-rays.

<sup>i</sup> RF/anti-CCP Ab indicates rheumatoid factor or anti-cyclic citrullinated peptide antibody.
<table>
<thead>
<tr>
<th>EVENT</th>
<th>Visit window (Days)</th>
<th>Screening</th>
<th>Day 1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>W2</th>
<th>W4</th>
<th>W8</th>
<th>W12</th>
<th>W14</th>
<th>W16</th>
<th>W20</th>
<th>W24</th>
<th>W26</th>
<th>W30</th>
<th>W36</th>
<th>W44</th>
<th>W52/ET&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Follow Up&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Post Treatment Week 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subject’s global assessment&lt;sup&gt;l&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HAQ-DI and Pain Scale&lt;sup&gt;l&lt;/sup&gt;</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FACIT-Fatigue, SF-36, WPAI-RA, EQ-5D&lt;sup&gt;l&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthcare resource utilization questionnaire and TSQM&lt;sup&gt;l&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exploratory patient reported outcomes&lt;sup&gt;l&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physician’s global assessment</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vital signs&lt;sup&gt;j&lt;/sup&gt;</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Symptom-driven physical examination</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12-Lead ECG</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SJ C66/TJC68</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weight</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVENT</td>
<td>Treatment Period</td>
<td>Follow Up$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit window (Days)</td>
<td>Day 1$^b$</td>
<td>W2</td>
<td>W4</td>
<td>W8</td>
<td>W12</td>
<td>W14</td>
<td>W16</td>
<td>W20</td>
<td>W24</td>
<td>W26</td>
<td>W30</td>
<td>W36</td>
<td>W44</td>
<td>W52/ET$^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visits</td>
<td>Screening $^a$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit window (Days)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy test $^b$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hematology and Chemistry $^1$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum CRP</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinalysis and urine for biomarkers$^m$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantitative Ig</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrine: FSH, TSH, and HbA1c</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipid profile (fasting) $^1$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genomic substudy $^n$ (optional)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PK blood samples $^o$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker Blood Samples: Plasma biomarker, serum biomarker, transcriptome, TCR/BCR repertoire $^o$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vIPBMC &amp; Leukocyte subsets$^q$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serology $^l$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral Monitoring $^s$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory lab panel $^l$</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CONFIDENTIAL
### Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

<table>
<thead>
<tr>
<th>EVENT</th>
<th>Treatment Period</th>
<th>Follow Up*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day 1(^b)</td>
<td>W2</td>
</tr>
<tr>
<td>Visits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit window (Days)</td>
<td>Screening(^a)</td>
<td>X</td>
</tr>
<tr>
<td>Imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand, wrist and feet radiographs(^d)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>MRI substudy (optional)(^e)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Carotid Artery Ultrasound(^f)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Assessment of Adverse Events</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Randomization/reassignment</td>
<td></td>
<td>X(^x)</td>
</tr>
<tr>
<td>Study Drug(s) Dispensation(^b)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Self-injection (or caregiver) training(^cc)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Review entry criteria for LTE (if applicable)</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

\(^a\) Screening window is 28 days prior to Day 1, the window may be extended to up to 42 days prior to the Day 1 visit for subjects who require repeat collection of radiographs.

\(^b\) Subject will begin medication on Day 1, following baseline assessments and randomization.

\(^c\) Early Termination Visit to be performed if subject discontinues before Week 52.

\(^d\) The follow-up visit will be performed only for subjects discontinuing prematurely from the study and for those completing Week 52, but not entering the LTE.

\(^e\) Eligibility criteria check based on the laboratory results from the Screening visit, TJC/SJC at screening and Day 1 and for female subjects of childbearing potential, the urine pregnancy test at Day 1.

\(^f\) Medical History includes smoking status, average weekly alcohol consumption, family history of coronary heart disease (CHD) any other chronic medical conditions or prior surgeries.

\(^g\) At Screening, Week 24, and Week 52 (or at ET), a complete physical examination should be performed. Symptom-directed exams should be performed at all other visits.

\(^h\) Height will be measured at Screening only.

\(^x\) Chest X-ray should be performed unless performed in the previous 3 months, and results are available at the site.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

i Subject reported outcomes are recommended to be performed at the beginning of each visit prior to any other visit-related procedures (other than signing of informed consent). Subject reported outcomes should be performed where available.

j Vital signs are defined as heart rate, respiratory rate, blood pressure (systolic and diastolic) and body temperature.

k For female subjects of childbearing potential only. To be performed on a serum sample at Screening and on urine samples for other visits. For female subjects entering the LTE, at the Week 52 visit pregnancy test should be performed on serum. During the periods where visits are every 6-8 weeks, women should continue to have pregnancy tests every 4 weeks, using home pregnancy urine tests, that will be provided to them. The site will call the subject every 4 weeks to obtain results of these pregnancy tests and will record the information in the source documents and CRF. If any pregnancy test is positive, study drug should be immediately interrupted and should come to the site for serum pregnancy test.

l Refer to list provided in Laboratory assessment table (Appendix 6).

m Urine for biomarkers is to be collected at Day 1 and Weeks 12, 24, 52 and ET (if applicable)

n Only for subjects who consent to the optional genomic substudy. The sample can be obtained at a subsequent visit if not obtained at the Day 1 visit.

o The PK sample at Week 4 should be collected at least 30 minutes post study drug dose in the clinic. PK samples at Week 12 and Week 24 should be collected prior to study drug dose (within 2 hours prior to dosing). The PK sample at Week 52 or ET can be collected at any time during the visit.

p Subjects that consent to the optional PK substudy will have additional PK samples collected prior to dose in the clinic (between Week 2 and Week 8, inclusive) and at 0.5, 1, 2, 3, 4, and 6 hours post dose (collection window: ±15 minutes for the 0.5 and 1 hour PK sample; ±30 minutes for the other time points). Subjects who are no longer on study drugs, are excluded from the optional PK substudy and no PK samples are collected.

q Details of biomarker collection are outlined in lab manual

r For sites within US and Canada only

s Viral monitoring for HBV or HCV as applicable (see Exclusion criteria, Section 4.3)

t At sites participating in carotid artery ultrasounds only. The following blood samples should be drawn in a fasting state; HbA1c, Leptin, LDL particle, Homocysteine, Apo A1/B

u Radiographs performed after enrollment may be done 67 days of the scheduled visit. If subject discontinues study early and has had X-rays <12 weeks prior to the ET visit, X-rays do not need to be repeated at the ET visit

v MRI to be performed within +7 days after randomization on Day 1 visit

w When available, at selected sites only

x Upon qualification for the study

y At Week 14, subjects who failed to achieve both a TJC and SJC improvement of 420% will be treated as inadequate responders and discontinue study drugs but will continue with study visits and assessments

z Subjects on placebo will be switched from placebo to filgotinib at Week 24

aa Subjects failing to maintain at least a 20% improvement from Day 1 in TJC and SJC confirmed at 2 consecutive visits will discontinue study drugs but will continue with study visits and assessments

bb Dispensation of study drug(s) according to IWRS manual

cc Self-injection or caregiver training will occur at Day 1. This can be repeated at Week 2 as required.
Appendix 3. Management of Clinical and Laboratory Adverse Events

<table>
<thead>
<tr>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>May continue dosing at the discretion of the investigator</td>
<td>Repeat lab to confirm toxicity grade</td>
<td>Repeat lab to confirm toxicity grade</td>
<td>If confirmed and possibly or probably related to investigational medicinal products, discontinue investigational medicinal products dosing permanently and follow at periodic intervals at least weekly until a return to baseline or is otherwise explained</td>
</tr>
</tbody>
</table>

If confirmed and possibly and/or probably related to investigational medicinal products:

1. Withhold investigational medicinal products until ≤ Grade 2
2. Restart all investigational medicinal products at full dose

If Grade 3 or 4 recurrence that is confirmed and possibly or probably related to investigational medicinal products, discontinue all investigational medicinal products dosing permanently

If Grade 3 or 4 recurrence that is considered unrelated to investigational medicinal products, continue all investigational medicinal products at the same dose at the discretion of the investigator
Appendix 4. Common Terminology Criteria for Adverse Events (CTCAE) v4.03

CTCAE v4.03 can be accessed from the below link:


The only modification to the CTCAE criteria is the addition of a Grade 1 upper respiratory infection as follows:

<table>
<thead>
<tr>
<th>CTCAE v4.0 Term</th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>CTCAE v4.03 AE Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper respiratory infection</td>
<td>Mild symptoms; symptomatic relief (eg, cough suppressant, decongestant)</td>
<td>Moderate symptoms; oral intervention indicated (eg, antibiotic, antifungal, antiviral)</td>
<td>IV antibiotic, antifungal, or antiviral intervention indicated; radiologic, endoscopic, or operative intervention indicated</td>
<td>Life-threatening consequences; urgent intervention indicated</td>
<td>Death</td>
<td></td>
<td>A disorder characterized by an infectious process involving the upper respiratory tract (nose, paranasal sinuses, pharynx, larynx, or trachea).</td>
</tr>
</tbody>
</table>
Appendix 5. Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Methods

The administration of filgotinib in embryo-fetal animal development studies resulted in decreased numbers of viable rat fetuses, increased resorptions, and visceral and skeletal malformations. Similar effects were noted in the rabbit. A safety margin relative to human exposure has not been identified. Pregnancy is contraindicated during use of filgotinib.

For participation in this study, all subjects of childbearing potential must agree to the use of highly effective contraception as outlined below.

1) Definitions

a) Definition of Childbearing Potential

For the purposes of this study, a female-born subject is considered of childbearing potential following the initiation of puberty (Tanner stage 2) until becoming post-menopausal, unless permanently sterile or with medically documented ovarian failure.

Women are considered to be in a postmenopausal state when they are ≥ 54 years of age with cessation of previously occurring menses for ≥ 12 months without an alternative cause. In addition, women of any age with amenorrhea ≥ 12 months may also be considered postmenopausal if their FSH level is in the postmenopausal range at screening and they are not using hormonal contraception or hormonal replacement therapy.

Permanent sterilization includes hysterectomy, bilateral oophorectomy, or bilateral salpingectomy in a female subject of any age.

b) Definition of Male Fertility

For the purposes of this study, a male-born subject is considered fertile after the initiation of puberty unless permanently sterile by bilateral orchidectomy or has medical documentation of permanent male infertility.

2) Contraception for Female Subjects

a) Study Drug Effects on Pregnancy and Hormonal Contraception

Filgotinib is contraindicated in pregnancy as there is a possibility of human teratogenicity/fetotoxicity in early pregnancy based on non-clinical data. A dedicated study assessing the impact of filgotinib on the efficacy of hormonally-based contraceptives (with ovulation inhibition as mechanism of action) has not yet been performed to fully verify the absence of any clinically significant interaction between filgotinib and oral contraceptives.

However, clinically relevant drug-drug interactions between hormonally-based contraceptives and filgotinib or its active metabolite GS-829845 are not expected. In vitro studies in HepaRG cells using reference substrates indicate that filgotinib and GS-829845 do not induce CYP1A2,
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

Final
Amendment 1

CY P2B6, or CY P3A4, enzymes involved in metabolism of common hormonal contraceptives. Preliminary results using primary human hepatocytes also confirmed that filgotinib and GS-829845 did not induce mRNA expression of CY P1A2, CY P2B6, or CY P3A4. The lack of induction potential of filgotinib and GS-829845 on CY P3A4 was further confirmed in a clinical drug-drug interaction study, where multiple dose administration of 200 mg filgotinib did not affect the pharmacokinetics of midazolam (a prototypical in vivo probe CY P3A4 substrate). Additionally, clinically relevant inhibition of CY Ps or UGTs during filgotinib treatment is unlikely, based on in vitro data.

Based on the totality of the in vitro and clinical data, clinically relevant drug interactions between filgotinib or GS-829845 with hormonal contraceptives are not expected. For female subjects, hormonal contraceptives will be permitted as a form of contraception when used in conjunction with a barrier method (preferably male condom). For male subjects, male condom should be used; for their female partners of childbearing potential, an accepted contraceptive method should also be considered. Details are outlined below.

Please refer to the latest version of the filgotinib investigator’s brochure for additional information.

b) Contraception for Female Subjects of Childbearing Potential

The inclusion of female subjects of childbearing potential requires the use of highly effective contraceptive measures. Women must have a negative serum pregnancy test at Screening and a negative urine pregnancy test on the Baseline/Day 1 visit prior to randomization. Pregnancy tests will be performed at monthly intervals thereafter. In the event of a delayed menstrual period (> one month between menstruations), a pregnancy test must be performed to rule out pregnancy. This is true even for women of childbearing potential with infrequent or irregular periods. Female subjects must agree to use one of the following methods from Screening until 35 days following the last dose of study drug.

- Complete abstinence from intercourse of reproductive potential. Abstinence is an acceptable method of contraception only when it is in line with the subject’s preferred and usual lifestyle.

  Or

- Consistent and correct use of 1 of the following methods of birth control listed below.
  - Intravaginal device (IUD) with a failure rate of <1% per year
  - Tubal sterilization
  - Essure micro-insert system (provided confirmation of success 3 months after procedure)
-- V asectomy in the male partner (provided that the partner is the sole sexual partner and had confirmation of surgical success at least 3 months after procedure, with documentation of sperm-free ejaculate)

- These above described methods are considered preferred methods of highly effective contraception in this protocol.

Female subjects who wish to use a hormonally based method must agree to use it in conjunction with a barrier method (used either by the female subject or by her male partner). Female subjects who utilize a hormonal contraceptive as one of their birth control methods must have consistently used the same method for at least three months prior to study dosing. Hormonally-based contraceptives and barrier methods permitted for use in this protocol are as follows:

- Hormonal methods (subject must agree to use with a barrier method, preferably, with a male condom)
  -- Oral contraceptives (either combined estrogen/progesterin or progesterone only)
  -- Injectable progesterone
  -- Implants of levonorgestrel
  -- Transdermal contraceptive patch
  -- Contraceptive vaginal ring

- Barrier methods (subject must agree to use with a hormonal method)
  -- Male or female condom, with or without spermicide
  -- Diaphragm with spermicide
  -- Cervical cap with spermicide
  -- Sponge with spermicide

- All female subjects must also agree to refrain from egg donation and in vitro fertilization during study participation and for at least 35 days after the last study drug dose.

3) Contraception for Male Subjects

It is theoretically possible that a relevant systemic concentration may be achieved in a female partner from exposure to the male subject’s seminal fluid. Therefore, male subjects with female partners of childbearing potential must agree to use condoms during study participation and for 90 days after the last study drug dose. Female partners of male study subjects should consider using one of the above methods of contraception as well. Male subjects must also agree to refrain from sperm donation during treatment and until at least 90 days after the end of dosing.
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

4) Unacceptable Birth Control Methods

Birth control methods that are unacceptable include periodic abstinence (e.g., calendar, ovulation, symptothermal, post-ovulation methods), withdrawal (coitus interruptus), spermicides only, and lactational amenorrhea method (LAM). Female condom and male condom should not be used together.

5) Procedures to be Followed in the Event of Pregnancy

Subjects will be instructed to notify the investigator if they become pregnant at any time during the study, or if they become pregnant within 35 days of last study drug dose. Subjects who become pregnant or who suspect that they are pregnant during the study must report the information to the investigator and discontinue study drug immediately. Subjects whose partner has become pregnant or suspects she is pregnant during the study are to report the information to the investigator.

Instructions for reporting pregnancy, partner pregnancy, and pregnancy outcome are outlined in Section 7.7.2.1.

Pregnancy Testing

All females of childbearing potential will have urine pregnancy testing every 4 weeks during the dosing period through 35 days after their last dose of study drug. During the periods where study visits are every 6-8 weeks, women should continue to have pregnancy tests every 4 weeks, using home pregnancy urine tests, that will be provided to them. The site will call the subject every 4 weeks to obtain results of these pregnancy tests and will record the information in the source documents and CRF. If a positive urine pregnancy test is reported, the subject will be asked to return to the clinic for a confirmatory serum pregnancy test.
# Appendix 6. Laboratory Assessment Table

<table>
<thead>
<tr>
<th>Hematology</th>
<th>Chemistry</th>
<th>Urinalysis</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematocrit</td>
<td>Alkaline phosphatase</td>
<td>Appearance</td>
<td>Urine drug screen for:</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>Aspartate aminotransferase (AST)</td>
<td>Blood</td>
<td>Amphetamines</td>
</tr>
<tr>
<td>Platelet count</td>
<td>Alanine aminotransferase (ALT)</td>
<td>Color</td>
<td>Cocaine</td>
</tr>
<tr>
<td>Red blood cell (RBC) count</td>
<td>Gamma-glutamyl transpeptidase (GGT)</td>
<td>Glucose</td>
<td>Methadone</td>
</tr>
<tr>
<td>Reb blood cell indices</td>
<td>Total bilirubin</td>
<td>Specific gravity</td>
<td>Opiates</td>
</tr>
<tr>
<td>White blood cell (WBC) count</td>
<td>Direct and indirect bilirubin</td>
<td>Nitrites</td>
<td>Leukocyte subsets*</td>
</tr>
<tr>
<td>Differentials (absolute and percentage), including:</td>
<td>Total protein</td>
<td>Leukocyte esterase</td>
<td>vfPBM*C</td>
</tr>
<tr>
<td>Leukocytes</td>
<td>Albumin</td>
<td>pH</td>
<td>C-reactive protein (hsCRP)</td>
</tr>
<tr>
<td>Monocytes</td>
<td>Bicarbonate</td>
<td>Protein</td>
<td>Rheumatoid factor and anticyclic citrullinated peptide antibody (RF/anti-CCP)</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>Blood urea nitrogen (BUN)</td>
<td>Urobilinogen</td>
<td>Quantiferon TB GOLD</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>Calcium</td>
<td>Reflex to microscopic urinalysis if dipstick result is abnormal.</td>
<td>Quantitative Ig subclasses at Screening, Week 24 and 52:</td>
</tr>
<tr>
<td>Basophils</td>
<td>Chloride</td>
<td></td>
<td>δ Leptin</td>
</tr>
<tr>
<td>Mean corpuscular volume (MCV)</td>
<td>CPK</td>
<td></td>
<td>δ LDL particle</td>
</tr>
<tr>
<td></td>
<td>Serum creatinine</td>
<td></td>
<td>δ Homocystein</td>
</tr>
<tr>
<td></td>
<td>Glucose</td>
<td></td>
<td>δ Apo A1/B</td>
</tr>
<tr>
<td></td>
<td>Phosphorus</td>
<td></td>
<td>δ Hb A1c</td>
</tr>
<tr>
<td></td>
<td>Magnesium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Potassium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sodium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrine at screening only</td>
<td>Amylase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemoglobin A1c</td>
<td>Lipase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSH</td>
<td>Uric Acid (screening)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSH (for women only)</td>
<td>Lipid profile (fasting):</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Triglycerides</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cholesterol and its subfractions (high-density lipoprotein [HDL] and low-density lipoprotein [LDL])</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serology</td>
<td>Hepatitis B, and core A b (if positive core Ab, then reflex Hep B DNA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy</td>
<td>Hepatitis C A b (if positive, then reflex HCV RNA)</td>
<td>In females of childbearing potential:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HIV</td>
<td>Serum β-hCG (Screening and if positive urine β-hCG)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urine β-hCG (all other visits)</td>
<td></td>
</tr>
</tbody>
</table>

* vfPBM*C and Leukocyte subsets US and Canada only.
* During the periods where study visits are every 6-8 weeks, women should continue to have pregnancy tests every 4 weeks, using home pregnancy urine tests, that will be provided to them. The site will call the subject every 4 weeks to obtain results of these pregnancy tests and will record the information in the source documents and CRF.
Appendix 7. List of Joints to be Evaluated (66/68 Joint Count)

An overview of the joints assessed is provided below (left and right):

- Temporomandibular
- Steroclavicular
- Acromioclavicular
- Shoulder
- Elbow
- Wrist
- Metacarpophalangeal: first, second, third, fourth, fifth
- Proximal interphalangeal: first, second, third, fourth, fifth
- Distal interphalangeal: second, third, fourth, fifth
- Hip
- Knee
- Ankle
- Tarsus
- Metatarsophalangeal: first, second, third, fourth, fifth
- Proximal interphalangeal: first, second, third, fourth, fifth

Replaced (or otherwise not assessable) joints should be documented at screening and omitted from further evaluation during the study.

---

2 Assessed for tenderness only
### Appendix 8. The 2010 American College of Rheumatology-European League Against Rheumatism Collaborative Initiative Classification Criteria for Rheumatoid Arthritis (Aletaha et al 2010)

<table>
<thead>
<tr>
<th>Target population (Who should be tested?)</th>
<th>Patients who have at least 1 joint with definite clinical synovitis (swelling) (^a) with the synovitis not better explained by another disease (^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification criteria for RA (score-based algorithm: add score of categories A - D; a score of (\geq 6/10) is needed for classification of a patient as having definite RA) (^c)</td>
<td></td>
</tr>
<tr>
<td>A. Joint involvement (^d)</td>
<td></td>
</tr>
<tr>
<td>1 large joint (^c)</td>
<td>0</td>
</tr>
<tr>
<td>2-10 large joints</td>
<td>1</td>
</tr>
<tr>
<td>1-3 small joints (with or without involvement of large joints) (^f)</td>
<td>2</td>
</tr>
<tr>
<td>4-10 small joints (with or without involvement of large joints)</td>
<td>3</td>
</tr>
<tr>
<td>&gt;10 joints (at least 1 small joint) (^g)</td>
<td>5</td>
</tr>
<tr>
<td>B. Serology (at least 1 test result is needed for classification) (^b)</td>
<td></td>
</tr>
<tr>
<td>Negative RF and negative ACPA</td>
<td>0</td>
</tr>
<tr>
<td>Low-positive RF or low-positive ACPA</td>
<td>2</td>
</tr>
<tr>
<td>High-positive RF or high-positive ACPA</td>
<td>3</td>
</tr>
<tr>
<td>C. Acute-phase reactants (at least 1 test result is needed for classification) (^i)</td>
<td></td>
</tr>
<tr>
<td>Normal CRP and normal ESR</td>
<td>0</td>
</tr>
<tr>
<td>Abnormal CRP or abnormal ESR</td>
<td>1</td>
</tr>
<tr>
<td>D. Duration of symptoms (^i)</td>
<td></td>
</tr>
<tr>
<td>&lt;6 weeks</td>
<td>0</td>
</tr>
<tr>
<td>(\geq 6) weeks</td>
<td>1</td>
</tr>
</tbody>
</table>

- The criteria are aimed at classification of newly presenting patients. In addition, patients with erosive disease typical of rheumatoid arthritis (RA) with a history compatible with prior fulfillment of the 2010 criteria should be classified as having RA. Patients with longstanding disease, including those whose disease is inactive (with or without treatment) who, based on retrospectively available data, have previously fulfilled the 2010 criteria should be classified as having RA.
- Differential diagnoses vary among patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted.
- Although patients with a score of \(<6/10\) are not classifiable as having RA, their status can be reassessed and the criteria might be fulfilled cumulatively over time.
- Joint involvement refers to any swollen or tender joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are excluded from assessment. Categories of joint distribution are classified according to the location and number of involved joints, with placement into the highest category possible based on the pattern of joint involvement.
- "Large joints" refers to shoulders, elbows, hips, knees, and ankles.
- "Small joints" refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.
- In this category, at least 1 of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (eg, temporomandibular, acromioclavicular, sternoclavicular).
Filgotinib
Protocol GS-US-417-0301
Gilead Sciences, Inc.

h  Negative refers to IU values that are less than or equal to the upper limit of normal (ULN) for the laboratory and assay; low-positive refers to IU values that are higher than the ULN but 1-6 times the ULN for the laboratory and assay; high-positive refers to IU values that are >3 times the ULN for the laboratory and assay. Where rheumatoid factor (RF) information is only available as positive or negative, a positive result should be scored as low-positive for RF.

ACP A = anti-citrullinated protein antibody.

i Normal/abnormal is determined by local laboratory standards. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.

j Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (e.g., pain, swelling, tenderness) of joints that are clinically involved at the time of assessment, regardless of treatment status.
Appendix 9. American College of Rheumatology Response Evaluations/ Preliminary Definition of Improvement in Rheumatoid Arthritis (Felson et al 1995)

ACR 20 Required

≥20% improvement in tender joint count, AND ≥20% improvement in swollen joint count, AND ≥20% improvement in at least 3 of the following 5 items:

- Patient pain assessment
- Patient global assessment of disease activity
- Physician global assessment of disease activity
- Patient’s assessment of physical function (HAQ-DI)
- A cute-phase reactant (CRP)

The following lists the disease activity measure followed by the method of assessment

1. Tender joint count

ACR tender joint count is an assessment of 68. The joint count should be done by scoring several different aspects of tenderness, as assessed by pressure and joint manipulation on physical examination. The information on various types of tenderness should then be collapsed into a single tender-versus-nontender dichotomy.

2. Swollen joint count

ACR swollen joint count is an assessment of 66. Joints are classified as either swollen or not swollen.

3. Patient’s assessment of pain

The pain score from the HAQ-DI will be used to calculate ACR response.
4. Patient’s global assessment of disease activity

A horizontal, visual analog scale will be used to provide the patient’s overall assessment of how the arthritis is doing.

Place a mark on the line below to indicate how you assess your current rheumatoid arthritis disease activity:

| No arthritis | Severe arthritis |

5. Physician’s global assessment of disease activity

A horizontal visual analog scale will be used to measure the physician’s assessment of the patient’s current disease activity.

Place a mark on the line below to indicate RA disease activity (independent of the subject’s self-assessment):

| No Disease Activity | Maximum Disease Activity |

6. Patient’s assessment of physical function

The HAQ-DI will be used to provide a patient’s self-assessment of physical function.

7. Acute-phase reactant value

C-reactive protein level as measured at the central laboratory
Appendix 10. Disease Activity Score (DAS28) { Prevoo et al 1995 }

Assessments of RA in patients by the Disease Activity Score (modified to include the 28 joint counts according to Smolen® 1995) will be conducted at the measured timepoints. The DAS28 consists of a composite score of the following variables: tender joint count, swollen joint count, CRP, and patient global score. The following equation will be used to calculate the DAS28-CRP

- DAS28(CRP) = 0.56 * TJ C28 + 0.28 * SJ C28 + 0.36 ln(CRP + 1) + 0.014 (patients global VAS) + 0.96

- TJ C28 = number of joints tender out of 28
- SJ C28 = number of joints swollen out of 28
- CRP = C-reactive protein
- Patient global VAS as defined in Appendix 9
Appendix 11. Procedures and Specifications

Complete Physical Examination

A complete physical examination should include source documentation of general appearance, and the following body systems: Head, neck and thyroid; eyes, ears, nose, throat, mouth and tongue; chest (excluding breasts); respiratory; cardiovascular; lymph nodes; abdomen; skin, hair, nails; musculoskeletal; and neurological.

Blood Pressure

Assessment of vital signs will include measurement of resting blood pressure, pulse, respiratory rate, and temperature.

Blood pressure will be measured using the following standardized process:

- Subject should be resting for ≥ 5 minutes with feet flat on the floor and measurement arm supported so that the midpoint of the manometer cuff is at heart level;
- Use a mercury sphygmomanometer or automatic blood pressure device with an appropriately sized cuff with the bladder centered over the brachial artery;
- Measure and record the blood pressure to the nearest 2 mmHg mark on the manometer or to the nearest whole number on an automatic device.

Creatinine Clearance

Creatinine clearance is calculated by the Cockcroft-Gault equation (Cockcroft et al. 1976) using actual body weight (BW).

\[
\text{Male: } \quad \text{CL}_c (\text{mL/min}) = \frac{[140 - \text{age (years)}] \times \text{BW(kg)}}{72 \times S_c} \\
\text{Female: } \quad \text{CL}_c (\text{mL/min}) = \frac{[140 - \text{age (years)}] \times \text{BW(kg)} \times 0.85}{72 \times S_c}
\]

\(S_c\) = serum creatinine (mg/dL)

12-Lead ECG

Subjects should be resting in a supine position for ≥ 5 minutes prior to making a recording.

The investigator (or qualified designee) should review the ECG traces recorded in real time for clinically significant abnormalities. On treatment ECGs should be compared to the subject’s Screening ECG as part of routine safety monitoring;
<table>
<thead>
<tr>
<th>Study Title:</th>
<th>A Randomized, Double-blind, Placebo- and Active-controlled, Multicenter, Phase 3 Study to Assess the Efficacy and Safety of Filgotinib Administered for 52 weeks in Combination with Methotrexate to Subjects with Moderately to Severely Active Rheumatoid Arthritis Who Have an Inadequate Response to Methotrexate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Test Drug:</td>
<td>Filgotinib</td>
</tr>
<tr>
<td>Study Number:</td>
<td>GS-US-417-0301</td>
</tr>
<tr>
<td>Protocol Version (Date):</td>
<td>Amendment 1: 05 July 2016</td>
</tr>
<tr>
<td>Analysis Type:</td>
<td>Week 24 Analysis and Final Analysis</td>
</tr>
<tr>
<td>Analysis Plan Version:</td>
<td>Version 1.0</td>
</tr>
<tr>
<td>Analysis Plan Date:</td>
<td>05 February 2019</td>
</tr>
<tr>
<td>Analysis Plan Author(s):</td>
<td>Lei Ye, PhD</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>2</td>
</tr>
<tr>
<td>LIST OF IN-TEXT TABLES</td>
<td>5</td>
</tr>
<tr>
<td>LIST OF IN-TEXT FIGURES</td>
<td>5</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>6</td>
</tr>
<tr>
<td>PHARMACOKINETIC ABBREVIATIONS</td>
<td>9</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>10</td>
</tr>
<tr>
<td>1.1 Study Objectives</td>
<td>10</td>
</tr>
<tr>
<td>1.2 Study Design</td>
<td>11</td>
</tr>
<tr>
<td>1.3 Sample Size and Power</td>
<td>12</td>
</tr>
<tr>
<td>2. TYPE OF PLANNED ANALYSIS</td>
<td>14</td>
</tr>
<tr>
<td>2.1 Data Monitoring Committee Analyses</td>
<td>14</td>
</tr>
<tr>
<td>2.2 Week 24 Analysis</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Final Analysis</td>
<td>14</td>
</tr>
<tr>
<td>3. GENERAL CONSIDERATIONS FOR DATA ANALYSES</td>
<td>15</td>
</tr>
<tr>
<td>3.1 Analysis Sets</td>
<td>15</td>
</tr>
<tr>
<td>3.1.1 All Randomized Analysis Set</td>
<td>15</td>
</tr>
<tr>
<td>3.1.2 Full Analysis Set</td>
<td>15</td>
</tr>
<tr>
<td>3.1.3 Per-Protocol Analysis Set</td>
<td>15</td>
</tr>
<tr>
<td>3.1.4 Safety Analysis Set</td>
<td>17</td>
</tr>
<tr>
<td>3.1.5 Pharmacokinetic Analysis Set</td>
<td>17</td>
</tr>
<tr>
<td>3.1.6 Pharmacokinetic Substudy Analysis Set</td>
<td>17</td>
</tr>
<tr>
<td>3.2 Subject Grouping</td>
<td>17</td>
</tr>
<tr>
<td>3.3 Strata and Covariates</td>
<td>18</td>
</tr>
<tr>
<td>3.4 Examination of Subject Subgroups</td>
<td>18</td>
</tr>
<tr>
<td>3.5 Multiple Comparisons</td>
<td>19</td>
</tr>
<tr>
<td>3.6 Missing Data and Outliers</td>
<td>20</td>
</tr>
<tr>
<td>3.6.1 Missing Data</td>
<td>20</td>
</tr>
<tr>
<td>3.6.2 Outliers</td>
<td>21</td>
</tr>
<tr>
<td>3.7 Data Handling Conventions and Transformations</td>
<td>21</td>
</tr>
<tr>
<td>3.8 Analysis Visit Windows</td>
<td>22</td>
</tr>
<tr>
<td>3.8.1 Definition of Study Day</td>
<td>22</td>
</tr>
<tr>
<td>3.8.2 Analysis Visit Windows</td>
<td>22</td>
</tr>
<tr>
<td>3.8.3 Selection of Non-Efficacy Data in the Event of Multiple Records in an Analysis Visit Window</td>
<td>29</td>
</tr>
<tr>
<td>4. SUBJECT DISPOSITION</td>
<td>30</td>
</tr>
<tr>
<td>4.1 Subject Enrollment and Disposition</td>
<td>30</td>
</tr>
<tr>
<td>4.2 Extent of Study Drug Exposure and Adherence</td>
<td>31</td>
</tr>
<tr>
<td>4.2.1 Duration of Exposure to Study Drug</td>
<td>31</td>
</tr>
<tr>
<td>4.2.2 Adherence to Study Drug</td>
<td>32</td>
</tr>
<tr>
<td>4.3 Protocol Deviations</td>
<td>34</td>
</tr>
<tr>
<td>5. BASELINE CHARACTERISTICS</td>
<td>35</td>
</tr>
<tr>
<td>5.1 Demographics and Other Baseline Characteristics</td>
<td>35</td>
</tr>
<tr>
<td>5.2 Baseline Disease Characteristics</td>
<td>35</td>
</tr>
<tr>
<td>5.3 Medical History</td>
<td>37</td>
</tr>
</tbody>
</table>
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis Version 1.0

6. EFFICACY ANALYSES ........................................................................................................... 38
   6.1. General Considerations ...................................................................................................... 38
       6.1.1. Tender/Swollen Joint Counts (TJC/SJC) ..................................................................... 41
       6.1.2. Global Assessment of Disease Activity ..................................................................... 42
       6.1.3. Health Assessment Questionnaire Disability Index (HAQ-DI) ................................. 43
   6.2. Primary Efficacy Endpoint ............................................................................................... 43
       6.2.1. Definition of the Primary Efficacy Endpoint ................................................................. 43
       6.2.2. Statistical Hypothesis for the Primary Efficacy Endpoint ........................................... 44
       6.2.3. Primary Analysis of the Primary Efficacy Endpoint ..................................................... 44
       6.2.4. Sensitivity Analysis of the Primary Endpoint .............................................................. 45
       6.2.5. Subgroup Analysis of the Primary Endpoint .............................................................. 46
   6.3. Key Secondary Efficacy Endpoints .................................................................................... 46
       6.3.1. Definition of Key Secondary Efficacy Endpoints ......................................................... 47
       6.3.2. Analysis Methods for Key Secondary Efficacy Endpoints .......................................... 48
       6.3.3. Sensitivity Analysis for Key Secondary Efficacy Endpoints ....................................... 49
       6.3.4. Subgroup Analyses for Key Secondary Efficacy Endpoints ....................................... 52
   6.4. Other Secondary Efficacy Endpoints ............................................................................... 52
       6.4.1. Definition of Other Secondary Efficacy Endpoints ..................................................... 53
       6.4.2. Analysis Methods for Other Secondary Efficacy Endpoints ....................................... 56
   6.5. Changes From Protocol-Specified Efficacy Analyses ..................................................... 60

7. SAFETY ANALYSES ............................................................................................................... 61
   7.1. Adverse Events and Deaths ............................................................................................ 61
       7.1.1. Adverse Event Dictionary ............................................................................................ 61
       7.1.2. Adverse Event Severity ............................................................................................... 61
       7.1.3. Relationship of Adverse Events to Study Drug ............................................................. 61
       7.1.4. Serious Adverse Events ............................................................................................... 61
       7.1.5. Treatment-Emergent Adverse Events ........................................................................ 61
       7.1.6. Summaries of Adverse Events and Deaths ................................................................. 62
       7.1.7. Adverse Events of Special Interest ............................................................................ 64
       7.1.8. Additional Analysis for Adverse Events ..................................................................... 65
   7.2. Laboratory Evaluations .................................................................................................... 66
       7.2.1. Summaries of Numeric Laboratory Results ................................................................. 66
       7.2.2. Graded Laboratory Value ........................................................................................... 68
       7.2.3. Laboratory Evaluations of Special Interest ................................................................. 70
   7.3. Body Weight and Vital Signs ............................................................................................ 71
   7.4. Prior and Concomitant Medications .................................................................................. 71
       7.4.1. Prior Medications ......................................................................................................... 72
       7.4.2. Concomitant Medications .......................................................................................... 72
   7.5. Electrocardiogram Results .............................................................................................. 73
       7.5.1. Investigator Electrocardiogram Assessment ................................................................. 73
   7.6. Other Safety Measures ..................................................................................................... 73
   7.7. Changes From Protocol-Specified Safety Analyses .......................................................... 73

8. PHARMACOKINETIC ANALYSES .......................................................................................... 74
   8.1. PK Analyses Related to Intensive PK Sampling ................................................................. 74
       8.1.1. Estimation of Pharmacokinetic Parameters ................................................................. 74
       8.1.2. Pharmacokinetic Parameters ....................................................................................... 74
       8.1.3. Statistical Analysis Methods ....................................................................................... 75

9. BIOMARKER ANALYSIS ........................................................................................................ 77

10. REFERENCES ......................................................................................................................... 78

11. SOFTWARE ............................................................................................................................ 79
Supplemental material

Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Version 1.0

12. SAP REVISION

13. APPENDIX

Appendix 1. Lists of RA Medications
Appendix 2. Health Assessment Questionnaire Disability Index (HAQ-DI)
Appendix 3. van der Heijde modified Total Sharp Score (mTSS)
Appendix 4. Treatment Satisfaction Questionnaire for Medication (TSQM)
Appendix 5. Corticosteroids
Appendix 6. SAS Programming for Tipping Point Analysis for Binary Endpoint
Appendix 7. SAS Programming for Tipping Point Analysis for Continuous Endpoint
Appendix 8. Sample SAS Code for Analysis of Binary Endpoints
LIST OF IN-TEXT TABLES

Table 3-1. Analysis Visit Windows for On-Treatment Joint Count Assessment, HAQ-DI, SGA, PGA, CRP, Exploratory PRO, and Vital Signs, Weight, and Safety Laboratory Data (Lipid Data Excluded) for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized ......................................................... 24

Table 3-2. Analysis Visit Windows for On-Treatment Joint Count Assessment, HAQ-DI, SGA, PGA, CRP, Exploratory PRO, and Vital Signs, Weight, and Safety Laboratory Data (Lipid Data Excluded) for Subjects who were Initially Randomized to Placebo and Re-randomized ............................................................ 25

Table 3-3. Analysis Visit Windows for On-Treatment Lipid Data for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized .................................................. 25

Table 3-4. Analysis Visit Windows for On-Treatment mTSS for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized .................................................. 26

Table 3-5. Analysis Visit Windows for On-Treatment Lipid Data for Subjects who were Initially Randomized to Placebo and Re-randomized .................................................. 26

Table 3-6. Analysis Visit Windows for On-Treatment mTSS data for Subjects who were Initially Randomized to Placebo and Re-randomized .................................................. 26

Table 3-7. Analysis Visit Windows for On-Treatment HRUQ,TSQM and ECG Data for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized .................................................. 27

Table 3-8. Analysis Visit Windows for On-Treatment HRUQ,TSQM and ECG Data for Subjects who were Initially Randomized to Placebo and Re-randomized .................................................. 27

Table 3-9. Analysis Visit Windows for On-Treatment SF-36, FACIT-Fatigue, EQ-5D, and WPAI-RA for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized .................................................. 28

Table 3-10. Analysis Visit Windows for On-Treatment SF-36, FACIT-Fatigue, EQ-5D, and WPAI-RA for Subjects who were Initially Randomized to Placebo and Re-randomized .................................................. 28

Table 6-1. Composition of the 28 Joints .................................................................................................................. 42

Table 6-2. EULAR Response Criteria .................................................................................................................. 54

Table 8-1. Study Treatments and Associated Analytes .......................................................................................... 74

Table 8-2. Pharmacokinetic Parameters for Each Analyte ...................................................................................... 75

LIST OF IN-TEXT FIGURES

Figure 1-1. Study Design ................................................................................................................................. 12
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab</td>
<td>antibody</td>
</tr>
<tr>
<td>ACR</td>
<td>American College of Rheumatology</td>
</tr>
<tr>
<td>ACR20/50/70</td>
<td>American College of Rheumatology 20/50/70% improvement</td>
</tr>
<tr>
<td>AE</td>
<td>adverse event</td>
</tr>
<tr>
<td>AESIs</td>
<td>adverse events of special interest</td>
</tr>
<tr>
<td>ALP</td>
<td>alkaline phosphatase</td>
</tr>
<tr>
<td>ALT</td>
<td>alanine aminotransferase</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>analysis of covariance</td>
</tr>
<tr>
<td>anti-CCP Ab</td>
<td>anti-Cyclic Citrullinated Peptide Antibody</td>
</tr>
<tr>
<td>AST</td>
<td>aspartate aminotransferase</td>
</tr>
<tr>
<td>ATC</td>
<td>anatomical therapeutic chemical drug class</td>
</tr>
<tr>
<td>bDMARD</td>
<td>biological disease modifying anti-rheumatic drug</td>
</tr>
<tr>
<td>BLQ</td>
<td>below the limit of quantitation</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>CBC</td>
<td>complete blood count</td>
</tr>
<tr>
<td>CCP</td>
<td>cyclic citrullinated peptide</td>
</tr>
<tr>
<td>CDAI</td>
<td>clinical disease activity index</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CPK</td>
<td>creatine phosphokinase</td>
</tr>
<tr>
<td>CRF</td>
<td>case report form</td>
</tr>
<tr>
<td>CSR</td>
<td>clinical study report</td>
</tr>
<tr>
<td>CTCAE</td>
<td>Common Toxicity Criteria for Adverse Events</td>
</tr>
<tr>
<td>DAS28</td>
<td>disease activity score for 28 joint count</td>
</tr>
<tr>
<td>DMC</td>
<td>data monitoring committee</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiogram</td>
</tr>
<tr>
<td>eCRF</td>
<td>electronic case report form</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>EuroQol 5 Dimensions</td>
</tr>
<tr>
<td>ET</td>
<td>early termination</td>
</tr>
<tr>
<td>EULAR</td>
<td>European League Against Rheumatism</td>
</tr>
<tr>
<td>FACIT-Fatigue</td>
<td>Functional Assessment of Chronic Illness Therapy-Fatigue</td>
</tr>
<tr>
<td>FAS</td>
<td>full analysis set</td>
</tr>
<tr>
<td>GI</td>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>Gilead</td>
<td>Gilead Sciences</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>health assessment questionnaire-disability index</td>
</tr>
<tr>
<td>HLG</td>
<td>high-level group term</td>
</tr>
<tr>
<td>HLT</td>
<td>high-level term</td>
</tr>
<tr>
<td>HRQoL</td>
<td>health-related quality of life</td>
</tr>
<tr>
<td>HRUQ</td>
<td>healthcare resource utilization questionnaire</td>
</tr>
</tbody>
</table>
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis  Version 1.0

hsCRP  high-sensitivity C-Reactive Protein
ID  identification
IWRS  interactive web response system
LLOQ  lower limit of quantitation
LLT  lower-level term
LOCF  last observation carried forward
LOQ  limit of quantitation
LS  least squares
LTE  long-term extension
MedDRA  Medical Dictionary for Regulatory Activities
MI  multiple imputation
MMRM  mixed model repeated measures
MNAR  missing not at random
MST  MedDRA search term
mTSS  van der Heijde modified Total Sharp Score
MTX  methotrexate
NRI  non-responder imputation
OC  observed case
PD  pharmacodynamics
PGA  physician’s global assessment
PK  pharmacokinetic
PP  per-protocol
PT  preferred term
PTM  placebo to match
PVE  pharmacovigilance & epidemiology department
Q1, Q3  first quartile, third quartile
q.d.  quaque die (each day)
RA  rheumatoid arthritis
RF  rheumatoid factor
SAE  serious adverse event
SAP  statistical analysis plan
SD  standard deviation
SDAI  simplified disease activity index
SDC  smallest detectable change
SE  standard error
SF-36  36-item short form survey
SGA  subject’s global assessment
SJC  swollen joint count
SMQ  standardized MedDRA query
SOC  system organ class
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAE</td>
<td>treatment-emergent adverse event</td>
</tr>
<tr>
<td>TFLs</td>
<td>tables, figures, and listings</td>
</tr>
<tr>
<td>TNF</td>
<td>tumor necrosis factor-alpha</td>
</tr>
<tr>
<td>TJC</td>
<td>tender joint count</td>
</tr>
<tr>
<td>TSQM</td>
<td>treatment satisfaction questionnaire for medication</td>
</tr>
<tr>
<td>ULN</td>
<td>upper limit of normal</td>
</tr>
<tr>
<td>VAS</td>
<td>visual analog scale</td>
</tr>
<tr>
<td>VR</td>
<td>ventricular rate</td>
</tr>
<tr>
<td>WBC</td>
<td>white blood cell</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WPAI</td>
<td>Work Productivity and Activity Impairment</td>
</tr>
</tbody>
</table>
**PHARMACOKINETIC ABBREVIATIONS**

- **AUC<sub>last</sub>**  
  area under the concentration versus time curve from time zero to the last quantifiable concentration

- **AUC<sub>tau</sub>**  
  area under the concentration versus time curve over the dosing interval

- **C<sub>last</sub>**  
  last observed quantifiable concentration of the drug

- **C<sub>max</sub>**  
  maximum observed concentration of drug

- **C<sub>tau</sub>**  
  observed drug concentration at the end of the dosing interval

- **CL<sub>ss/F</sub>**  
  apparent oral clearance after administration of the drug;
  at steady state: \( \text{CL}<sub>ss/F</sub> = \frac{\text{Dose}}{\text{AUC}<sub>tau</sub>}, \) where \( \text{Dose} \) is the dose of the drug

- **t<sub>1/2</sub>**  
  estimate of the terminal elimination half-life of the drug, calculated by dividing the natural log of 2 by the terminal elimination rate constant \( (\lambda_2) \)

- **T<sub>last</sub>**  
  time (observed time point) of \( C<sub>last</sub> \)

- **T<sub>max</sub>**  
  time (observed time point) of \( C<sub>max</sub> \)

- **V<sub>D/F</sub>**  
  apparent volume of distribution of the drug

- **\( \lambda_2 \)**  
  terminal elimination rate constant, estimated by linear regression of the terminal elimination phase of the concentration of drug versus time curve
1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in tables, figures, and listings (TFLs) in the clinical study report (CSR) for Study GS-US-417-0301. This SAP is based on the study protocol Amendment 1 dated 05 July 2016 and the electronic case report form (eCRF). The SAP will be finalized before the unblinding of Week 24 analysis. Any changes made after the finalization of the SAP will be documented in the CSR.

1.1. Study Objectives

The primary objective of this study is as follows:

- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of rheumatoid arthritis (RA) as measured by the proportion of subjects achieving an American College of Rheumatology 20% improvement response (ACR20) at Week 12

The secondary objectives of this study are as follows:

- To evaluate the effects of filgotinib versus placebo as measured by the proportion of subjects achieving Disease Activity Score for 28 joint count using c-reactive protein (DAS28[CRP]) H3.2 at Week 12
- To evaluate the effect of filgotinib versus placebo on physical function as measured by change from Baseline in Health Assessment Questionnaire Disability Index (HAQ-DI) score at Week 12
- To evaluate the effects of filgotinib versus placebo for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28(CRP) < 2.6 at Week 24
- To evaluate the effects of filgotinib versus placebo on preservation of joint structure as measured by change from Baseline in van der Heijde modified Total Sharp Score (mTSS) at Week 24
- To evaluate the effects of filgotinib versus adalimumab for the treatment of signs and symptoms of RA as measured by the proportion of subjects achieving DAS28(CRP) H3.2 at Week 12
- To evaluate the safety and tolerability of filgotinib
- To evaluate the effects of filgotinib on work productivity, fatigue, and general quality of life as measured by 36-Item Short Form Survey (SF-36), Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue), EuroQol 5 Dimensions (EQ-5D) and Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA)
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

The exploratory objectives of this study are as follows:

- To characterize the pharmacokinetics (PK) of filgotinib and its metabolite (GS-829845, formerly G254445)
- To characterize the association of host genetics and other markers with disease severity, disease progression and treatment response to filgotinib in subjects with RA
- To evaluate the effects of filgotinib on healthcare resource utilization and other patient reported outcomes

1.2. Study Design

This is a randomized, double-blind, placebo- and active-controlled, Phase 3 study in adult male and female subjects with active RA who have an inadequate response to MTX (MTX-IR). The study is designed to evaluate the efficacy, safety, and tolerability of filgotinib as well as its effect on work productivity, fatigue, and quality of life.

Approximately 1650 subjects will be randomized in a 3:3:2:3 ratio to filgotinib 200 mg, filgotinib 100 mg, active comparator (adalimumab), or placebo to match (PTM) administered for up to 52 weeks, all in the context of a weekly stable dose of MTX:

- Filgotinib 200 mg group: filgotinib (200 mg q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
- Filgotinib 100 mg group: filgotinib (100 mg q.d.) + PTM filgotinib 200 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)
- Active comparator group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + adalimumab (40 mg s.c. injection q2w) (N=300)
- Placebo control group: PTM filgotinib 200 mg (PTM q.d.) + PTM filgotinib 100 mg (PTM q.d.) + PTM adalimumab (PTM s.c. injection q2w) (N=450)

Randomization will be stratified by geographic region, prior exposure to biologic disease modifying antirheumatic drugs (bDMARDs), and presence of rheumatoid factor (RF) or anti-CCP (cyclic citrullinated peptide) antibody (Ab) at screening.

At Week 14, subjects who have not achieved at least 20% improvement from Baseline in both swollen joint count (SJ C) and tender joint count (TJC) will discontinue investigational study drug dosing but will continue with study visits and assessments per protocol. All subjects meeting this criterion who discontinue from investigational therapy are to receive standard of care treatment for their RA as determined by the investigator.

At Week 24, all subjects assigned to placebo + MTX will be reassigned 1:1 to either filgotinib 100 mg q.d. or 200 mg q.d. in addition to MTX in a blinded fashion and will continue in the study per protocol up to Week 52.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

All subjects who continue on study drug will be evaluated for loss of therapeutic response from Week 30 through Week 52. Subjects failing to maintain at least a 20% improvement from Baseline in TJC and SJC, (which is confirmed at 2 consecutive visits), will discontinue from investigational study drug therapy but will continue with study visits and assessments per protocol. All subjects meeting this criterion who discontinue from investigational study drug dosing are to receive standard of care treatment for their RA as determined by the investigator.

All subjects who have received at least one dose of study drug and exit the study early will complete an early termination (ET) visit at the time of study discontinuation, with a follow up visit four weeks after the last dose of study drug (Post Treatment Week 4), regardless of dosing duration.

At completion of the 52-week dosing period, subjects who have not discontinued assigned study drug dosing, will be provided the option to enroll into a separate Long Term Extension (LTE) study (GS-US-417-0304).

Figure 1-1. Study Design

1.3. Sample Size and Power

Sample size is determined based on the superiority test of filgotinib 200 mg compared to placebo based on the change from Baseline in mTSS at Week 24. When assuming a difference of 0.4 between filgotinib and placebo on change from Baseline in mTSS at Week 24 and a common standard deviation of 1.85, 450 subjects in each of the filgotinib 200 mg group and placebo control group are required to obtain 90% power at a 2-sided 0.05-level.

A sample size of 450 subjects in each of the filgotinib groups and placebo control group will provide over 95% power to detect an increase in ACR20 response rate of 45% to 65% between the placebo control group and the filgotinib group respectively, using a 2-sided 0.05-level test.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

Version 1.0

Based on {Liu 2014}, 450 subjects in each of the filgotinib 200 mg group and placebo group, and 300 subjects in the adalimumab group, will provide over 90% power at 2-sided 0.05 significance level to demonstrate that filgotinib 200 mg preserves more than 50% of the effect of adalimumab with respect to the response rate of DAS28(CRP) HB.2 at Week 12, assuming both filgotinib 200 mg and adalimumab groups have similar response rates of DAS28(CRP) HB.2. Given this study has a placebo group, assay sensitivity can be established through a direct comparison of adalimumab to placebo.

In summary, the total planned sample size will be 1650 (450 for the filgotinib 200 mg, filgotinib 100 mg, placebo groups, and 300 for the active comparator group).
2. TYPE OF PLANNED ANALYSIS

2.1. Data Monitoring Committee Analyses

An external multidisciplinary Data Monitoring Committee (DMC) will review the progress of the study and perform interim reviews of safety data in order to protect subject welfare and preserve study integrity. To ensure the best interests of the participants, the DMC will recommend to the sponsor if the nature, frequency, and severity of adverse effects associated with the study treatment warrant the early termination of the study, the continuation of the study, or the continuation of the study with modifications.

The initial review has been conducted after approximately 100 subjects have completed at least 12 weeks of dosing in any one of the three RA phase 3 studies (GS-US-417-0301, GS-US-417-0302, and GS-US-417-0303). Regular DMC review of safety data will be scheduled approximately every 6 months following the initial data review meeting including all 3 RA phase 3 studies and a long-term extension study (GS-US-417-0304).

The DMC’s role and responsibilities and the scope of analysis to be provided to the DMC are provided in a mutually agreed upon charter, which defines the DMC membership, meeting logistics, and meeting frequency.

2.2. Week 24 Analysis

A planned Week 24 analysis will be conducted after all subjects have either completed their Week 24 visit or prematurely discontinued from the study. A pre-specified sponsor team including members of which are not actively involved in the conduct of the study, will review the Week 24 unblinded safety and efficacy analysis results. A memo and a list will be maintained which documents the individuals who have been granted access to the Week 24 unblinded results along with the justification for unblinding in accordance with Gilead SOPs. A data integrity and communication plan for the pre-specified Week 24 analysis will be finalized prior to unblinding. The Study Management Team members with direct involvement in the conduct of the study will either be replaced if they need to be unblinded, or remain blinded to treatment assignments throughout the trial, until all subjects have completed the planned study Week 52 visits and the database has been locked.

2.3. Final Analysis

After all subjects have completed the study, outstanding data queries have been resolved or adjudicated as unresolvable, and the data have been cleaned and finalized, the study blind will be broken and the final analysis of the data will be performed.
3. **GENERAL CONSIDERATIONS FOR DATA ANALYSES**

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of subjects in each category will be presented; for continuous variables, the number of subjects (n), mean, standard deviation (SD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

All statistical tests will be 2-sided and performed at the 5% significance level unless otherwise specified.

By-subject listings will be presented for all subjects in the All Randomized Analysis Set and sorted by subject identification (ID) number, visit date, and time (if applicable). Data collected on log forms, such as AEs, will be presented in chronological order within the subject. The treatment group to which subjects were randomized will be included in the listings, as well as age, sex at birth, race, and ethnicity.

3.1. **Analysis Sets**

Analysis sets define the subjects to be included in an analysis. Analysis sets and their definitions are provided in this section. Subjects included in each analysis set will be determined before the study blind is broken for analysis. The analysis set will be identified and included as a subtitle of each TFL.

For each analysis set, the number and percentage of subjects eligible for inclusion, as well as the number and percentage of subjects who were excluded and the reasons for their exclusion, will be summarized by treatment group.

A listing of reasons for exclusion from analysis sets will be provided by subject.

3.1.1. **All Randomized Analysis Set**

All Randomized Analysis Set includes all subjects who are randomized in the study. This is the primary analysis set for by-subject listings.

3.1.2. **Full Analysis Set**

The Full Analysis Set (FAS) includes all randomized subjects who received at least 1 dose of study drug. The study drugs in this study are filgotinib 200 mg, filgotinib 100 mg, adalimumab and PTMs. This is the primary analysis set for efficacy analyses.

3.1.3. **Per-Protocol Analysis Set**

The Per-Protocol (PP) Analysis Set includes subjects in the FAS who did not have major protocol deviations occurring prior to or on Week 12 that will affect the efficacy analysis, and were compliant (with on-treatment adherence ≥ 80% in the first 12-week period) to study treatment.
Major protocol deviations include the following:

- Violation of at least one of the following key eligibility criteria:
  1) Have a diagnosis of RA (2010 ACR/EULAR criteria for RA), and are ACR functional class I-III
  2) Have 6 swollen joints (from swollen joint count based on 66 joints [SJ C66]) and 6 tender joints (from tender joint count based on 68 joints [TJ C68]) at both Screening and Day 1 (need not be the same joints)
  3) Must meet at least one of the following parameters at Screening:
     a) ≥1 documented joint erosion on radiographs of the hands, wrists or feet by central reading AND a positive result for anti-CCP or RF (based on central laboratory)
        OR
     b) ≥3 documented erosions on radiographs of the hands, wrists or feet by central reading if both antibodies (ie, RF, anti-CCP) are negative (based on central laboratory)
        OR
     c) Serum CRP ≥ 6 mg/L (based on central laboratory)
  4) Ongoing treatment with a stable dose of MTX as described in protocol
  5) Subjects that have failed prior therapy with a bDMARD are not eligible to participate. Subjects with prior exposure to one bDMARD may be enrolled if there is documented evidence of limited exposure (ie, less than 3 months) to the bDMARD

- Received prohibited concomitant medications
- Background medications were changed in violation of protocol
- Improper study medication administration and unintended medication overdose
- Data were questionable because of significant site quality or compliance issues

Qualifications and identification of the specific major protocol deviations that result in exclusion from the PP Analysis Set will be determined while the study remains blinded, prior to database finalization.

The PP Analysis Set is the secondary analysis set for efficacy analyses on primary and key secondary endpoints at Week 12. Statistical analyses using PP Analysis Set are specified in Section 6.
3.1.4. Safety Analysis Set

The Safety Analysis Set includes all subjects who received at least 1 dose of study drug. This is the primary analysis set for safety analyses.

3.1.5. Pharmacokinetic Analysis Set

The Pharmacokinetic (PK) analysis set includes all subjects in the Safety Analysis Set who have at least 1 nonmissing plasma concentration data for filgotinib and/or its metabolite GS-829845. This is the primary analysis set for general PK analyses.

3.1.6. Pharmacokinetic Substudy Analysis Set

The PK Substudy Analysis Set includes all subjects in the Safety Analysis Set who enrolled into the PK Substudy, and have intensive concentration data to provide interpretable results for the specific parameters of interest for the analyte under evaluation. This is the primary analysis set for intensive PK analyses.

3.2. Subject Grouping

For analyses based on the All Randomized Analysis Set and FAS, subjects will be grouped according to the treatment to which they were randomized. For analyses based on the Safety Analysis Set and PP Analysis Set, subjects will be grouped according to actual treatment received. The actual treatment received will differ from the randomized treatment only when their actual treatment differs from randomized treatment for the entire treatment duration of each treatment period.

The treatment groups during the placebo controlled period are:

- Filgotinib 200 mg q.d.
- Filgotinib 100 mg q.d.
- Adalimumab
- Placebo

The treatment groups during the re-randomized period are:

- Filgotinib 200 mg q.d.
- Filgotinib 100 mg q.d.
- Adalimumab
- Placebo switch to filgotinib 200 mg q.d.
- Placebo switch to filgotinib 100 mg q.d.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan – Week 24 Analysis and Final Analysis     Version 1.0

For the PK Analysis Set, subjects will be grouped according to the actual treatment they received.

3.3. Strata and Covariates

Subjects will be randomly assigned to treatment groups via the interactive web response system (IWRS) in a 3:3:2:3 ratio to filgotinib 200 mg, filgotinib 100 mg, adalimumab or PTM using a stratified randomization schedule. Stratification will be based on the following variables:

- Geographic region (Group A) includes the following countries: Australia, Belgium, Canada, France, Germany, Ireland, Israel, Italy, Netherlands, New Zealand, Republic of Korea, South Africa, Singapore, Spain, Sweden, Switzerland, United Kingdom, and USA; Group B includes the following countries: Bulgaria, Croatia, Czech Republic, Estonia, Georgia, India, Moldova, Romania, Russia, Slovakia, Ukraine, Serbia, Hungary, Latvia and Poland; Group C includes the following countries: Argentina, Brazil, Chile, Colombia, Peru, Puerto Rico and Mexico; Group D includes the following countries: China, Hong Kong, Malaysia, Philippines, Taiwan, Thailand and Vietnam; Group E includes Japan;

- Prior exposure to bDMARDs (Yes or No)

- Presence of RF or anti-CCP Ab (Yes or No) (Presence of RF =No if RF < 15 IU/mL; Presence of anti-CCP Ab = No if anti-CCP Quant < 17 U/mL)

If there are discrepancies in stratification factor values between the IWRS and the clinical database, the values recorded in the clinical database will be used for analyses.

For efficacy endpoints, stratification factors and the baseline value of the efficacy variable(s) will be included as covariates in the efficacy analysis model, as specified in Section 6.

3.4. Examination of Subject Subgroups

The primary and key secondary efficacy endpoints will be examined using the following subgroups (but not limited to the ones listed below):

- Age (on the first dosing date of any study drug, < 65 or ≥ 65)

- Sex at birth (male or female)

- Race

- Baseline weight (< 60 kg, ≥ 60 kg to < 100 kg or ≥ 100 kg)

- Geographic region (A, B, C, D or E)

- Prior exposure to bDMARDs (Yes or No)

- Presence of RF or anti-CCP Ab (Yes or No)
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Version 1.0

- Duration of RA diagnosis on the first dosing date of study drug (< 1 year, ≥ 1 year to < 5 years, ≥ 5 years to < 10 years or ≥ 10 years)
- Disease activity on the first dosing date of study drug (DAS28[CRP] ≥ 5.1 or DAS28[CRP] > 5.1)
- Concurrent use of oral corticosteroids on the first dosing date of study drug (Yes or No)
- High-Sensitivity C-reactive Protein (hsCRP) (≥ 4 mg/L or < 4 mg/L)

Lists of bDMARDs and oral corticosteroids are provided in Appendix 1.

3.5. Multiple Comparisons

The primary endpoint for the study is the proportion of subjects who achieve an ACR20 response at Week 12. The primary analyses will consist of a superiority test of filgotinib 200 mg compared to placebo based on the ACR20 response rate at Week 12.

The following hypothesis testing for secondary analyses will commence after the primary analysis reaches statistical significance, and will be tested according to the hierarchical testing principle at the 2-sided 0.05 level. If a null hypothesis is not rejected, formal sequential testing will be stopped and only nominal significance will be reported for the remaining hypotheses:

1) Superiority of filgotinib 100 mg compared to placebo based on ACR20 response rate at Week 12

2) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in HAQ-DI at Week 12

3) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in HAQ-DI at Week 12

4) Superiority of filgotinib 200 mg compared to placebo based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

5) Superiority of filgotinib 100 mg compared to placebo based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

6) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in mTSS at Week 24

7) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in mTSS at Week 24

8) Non-inferiority of filgotinib 200 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) ≥ 3.2 at Week 12

CONFIDENTIAL

Page 19

05 February 2019
Filgotinib  
GS-US-417-0301  
Statistical Analysis Plan - Week 24 Analysis and Final Analysis  

Version 1.0

9) Non-inferiority of filgotinib 100 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) ≥ 3.2 at Week 12

10) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in SF-36 PCS at Week 12

11) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in SF-36 PCS at Week 12

12) Superiority of filgotinib 200 mg compared to placebo based on the change from Baseline in FACIT-Fatigue at Week 12

13) Superiority of filgotinib 100 mg compared to placebo based on the change from Baseline in FACIT-Fatigue at Week 12

14) Superiority of filgotinib 200 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) ≥ 3.2 at Week 12

15) Superiority of filgotinib 100 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

16) Non-inferiority of filgotinib 200 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

17) Non-inferiority of filgotinib 100 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

18) Superiority of filgotinib 200 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

19) Superiority of filgotinib 100 mg compared to adalimumab based on the proportion of subjects with DAS28(CRP) < 2.6 at Week 12

3.6. Missing Data and Outliers

3.6.1. Missing Data

In general, missing data will not be imputed unless methods for handling missing data are specified. Exceptions are presented in this document.

For missing last dosing date of study drug, imputation rules are described in Section 4.2.1. For partial date of initial RA diagnosis, imputation rules are described in Section 5.2. The handling of missing or incomplete dates for AE onset is described in Section 7.1.5.2, and for prior and concomitant medications in Section 7.4. Imputation rules adopted in the efficacy analyses are specified in Section 6.
3.6.2. Outliers

Outliers will be identified during the data management and data analysis process, but no sensitivity analyses will be conducted. All data will be included in the data analyses.

3.7. Data Handling Conventions and Transformations

In general, age (in years) on the date of the first dose of study drug will be used for analyses and presentation in listings. If an enrolled subject was not dosed with any study drug, the randomization date will be used instead of the first dosing date of study drug. For screen failures, the date the last informed consent was signed will be used for age calculation. If only birth year is collected on the CRF, '01 July' will be used for the unknown birth day and month for the purpose of age calculation. If only birth year and month are collected, '01_ will be used for the unknown birth day.

Non-PK Data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of `< x_` (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1, etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used for calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of `> x_` (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of `Hx_ or `h x_` (where x is considered the LOQ).

Natural logarithm transformation will be used for plasma/blood concentrations and analysis of PK parameters. Plasma concentration values that are below the limit of quantitation (BLQ) will be presented as `BLQ_` in the concentration data listing. Values that are BLQ will be treated as 0 at predose time points, and one-half the value of the LOQ at postbaseline time points.

The following conventions will be used for the presentation of summary and order statistics:

- If at least 1 subject has a concentration value of BLQ for the time point, the minimum value will be displayed as `BLQ_`

- If more than 25% of the subjects have a concentration data value of BLQ for a given time point, the minimum and Q1 values will be displayed as `BLQ_`

- If more than 50% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, and median values will be displayed as `BLQ_`
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

Version 1.0

- If more than 75% of the subjects have a concentration data value of BLQ for a given time point, the minimum, Q1, median, and Q3 values will be displayed as `BLQ`.

- If all subjects have concentration data values of BLQ for a given time point, all order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as `BLQ`.

PK parameters that are BLQ will be imputed as one-half LOQ before log transformation or statistical model fitting.

### 3.8. Analysis Visit Windows

#### 3.8.1. Definition of Study Day

The first dose date of individual study drug will be calculated separately for each study drug (ie, filgotinib 200mg, filgotinib 100mg, adalimumab and PTMs) in a treatment group. Study Day 1 is defined as the first dose date of any study drug, which is the minimum of the first dose dates of individual study drugs in a treatment group.

The last dose date of individual study drug will be calculated separately for each study drug in a treatment group. The last dose date for an individual study drug will be the end date on study drug administration CRF for the record where the `study drug was permanently withdrawn` flag is `Yes`. The last dose date of any study drug will be defined as the maximum of the last dose dates of individual study drugs in a treatment group.

Study Day will be calculated from the Study Day 1 and derived as follows:

- For postdose study days: Assessment Date ~ Study Day 1 + 1
- For days prior to the first dose: Assessment Date ~ Study Day 1

For subjects who were on placebo and then re-randomized, the On-Filgotinib Study Day 1 is defined as the first dose date of filgotinib which is identified by the first dispense date of filgotinib. The On-Filgotinib Study Day will be calculated from the On-Filgotinib Study Day 1 and derived as follows:

- For postdose study days of filgotinib: Assessment Date ~ On-Filgotinib Study Day 1 + 1
- For days prior to the first dose of filgotinib: Assessment Date ~ On-Filgotinib Study Day 1

#### 3.8.2. Analysis Visit Windows

Subject visits may not occur on protocol-specified days. Therefore, for the purpose of analysis, observations will be assigned to analysis windows.
In general, the baseline value will be the last nonmissing value on or prior to the first dose date of study drug, except for the mTSS baseline value. The mTSS baseline value will be derived based on the measurements collected on or prior to the first dose date of study drug + 14 days given the potential delay in scheduling an x-ray assessment. If multiple valid, nonmissing measurements exist in the analysis window as aforementioned, in general, the last nonmissing mTSS value on or prior to the first dose date of study drug will be selected as the baseline value; if all these multiple measurements occur at post first dosing, the value closest to the first dose date will be selected as the baseline value.

For subjects who were on placebo and then re-randomized, the baseline after Week 24 will be the last nonmissing value obtained on or prior to the first dose date of filgotinib, except for the mTSS baseline value. The mTSS baseline value will be derived based on the measurements collected on or prior to the first dose date of filgotinib + 14 days given the potential delay in scheduling an x-ray assessment. If multiple valid, nonmissing measurements exist in the analysis window as aforementioned, in general, the last nonmissing mTSS value on or prior to the first dose date of filgotinib will be selected as the baseline value; if all these multiple measurements occur at post first dosing of filgotinib, the value closest to the first dose date of filgotinib will be selected as the baseline value.

For efficacy endpoints, vital signs, ECG, weight, lipids and safety laboratory data, the analysis visit windows will be applied to data collected during the on-treatment period. The on-treatment period is defined as the last dose date of any study drug + 7 days.

The analysis windows for on-treatment efficacy endpoints including joint count assessment, HAQ-DI (including subject’s pain assessment), subject’s global assessment (SGA), physician’s global assessment (PGA), serum CRP, exploratory patient reported outcome (PRO), and vital signs, weight, and safety laboratory data (lipid data excluded) are provided in Table 3-1 and Table 3-2.
Table 3-1. Analysis Visit Windows for On-Treatment Joint Count Assessment, HAQ-DI, SGA, PGA, CRP, Exploratory PRO, and Vital Signs, Weight, and Safety Laboratory Data (Lipid Data Excluded) for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 2</td>
<td>15</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>Week 4</td>
<td>29</td>
<td>23</td>
<td>43</td>
</tr>
<tr>
<td>Week 8</td>
<td>57</td>
<td>44</td>
<td>71</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>72</td>
<td>92</td>
</tr>
<tr>
<td>Week 14</td>
<td>99</td>
<td>93</td>
<td>106</td>
</tr>
<tr>
<td>Week 16</td>
<td>113</td>
<td>107</td>
<td>127</td>
</tr>
<tr>
<td>Week 20</td>
<td>141</td>
<td>128</td>
<td>155</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>156</td>
<td>176</td>
</tr>
<tr>
<td>Week 26</td>
<td>183</td>
<td>177</td>
<td>197</td>
</tr>
<tr>
<td>Week 30</td>
<td>211</td>
<td>198</td>
<td>232</td>
</tr>
<tr>
<td>Week 36</td>
<td>253</td>
<td>233</td>
<td>281</td>
</tr>
<tr>
<td>Week 44</td>
<td>309</td>
<td>282</td>
<td>337</td>
</tr>
<tr>
<td>Week 52</td>
<td>365</td>
<td>338</td>
<td>h 365</td>
</tr>
</tbody>
</table>
**Table 3-2. Analysis Visit Windows for On-Treatment Joint Count Assessment, HAQ-DI, SGA, PGA, CRP, Exploratory PRO, and Vital Signs, Weight, and Safety Laboratory Data (Lipid Data Excluded) for Subjects who were Initially Randomized to Placebo and Re-randomized**

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 2</td>
<td>15</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>Week 4</td>
<td>29</td>
<td>23</td>
<td>43</td>
</tr>
<tr>
<td>Week 8</td>
<td>57</td>
<td>44</td>
<td>71</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>72</td>
<td>92</td>
</tr>
<tr>
<td>Week 14</td>
<td>99</td>
<td>93</td>
<td>106</td>
</tr>
<tr>
<td>Week 16</td>
<td>113</td>
<td>107</td>
<td>127</td>
</tr>
<tr>
<td>Week 20</td>
<td>141</td>
<td>128</td>
<td>155</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>156</td>
<td>First Dose Date of Filgotinib - Study Day 1 + 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>On-Filgotinib Study Day</th>
<th>On-Filgotinib Lower Limit</th>
<th>On-Filgotinib Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 26*</td>
<td>15</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Week 30*</td>
<td>43</td>
<td>30</td>
<td>64</td>
</tr>
<tr>
<td>Week 36*</td>
<td>85</td>
<td>65</td>
<td>113</td>
</tr>
<tr>
<td>Week 44*</td>
<td>141</td>
<td>114</td>
<td>169</td>
</tr>
<tr>
<td>Week 52*</td>
<td>197</td>
<td>170</td>
<td>h 197</td>
</tr>
</tbody>
</table>

* For subjects who were on placebo and re-randomized their analysis window after Week 24 will be assigned according to On-Filgotinib Study Day, On-Filgotinib Lower Limit and On-Filgotinib Upper Limit.

The analysis windows for on-treatment mTSS and lipid data are provided in Table 3-3, Table 3-4, Table 3-5, and Table 3-6.

**Table 3-3. Analysis Visit Windows for On-Treatment Lipid Data for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized**

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>2</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>267</td>
</tr>
<tr>
<td>Week 52</td>
<td>365</td>
<td>268</td>
<td>h 365</td>
</tr>
</tbody>
</table>
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Table 3-4. Analysis Visit Windows for On-Treatment mTSS for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>15</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>16</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>267</td>
</tr>
<tr>
<td>Week 52</td>
<td>365</td>
<td>268</td>
<td>h 365</td>
</tr>
</tbody>
</table>

Table 3-5. Analysis Visit Windows for On-Treatment Lipid Data for Subjects who were Initially Randomized to Placebo and Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>2</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>First Dose Date of Filgotinib - Study Day 1 + 1</td>
</tr>
<tr>
<td>Week 52*</td>
<td>On-Filgotinib Study Day</td>
<td>197</td>
<td>2</td>
</tr>
</tbody>
</table>

* For subjects who were on placebo and re-randomized their analysis window after Week 24 will be assigned according to On-Filgotinib Study Day, On-Filgotinib Lower Limit and On-Filgotinib Upper Limit.

Table 3-6. Analysis Visit Windows for On-Treatment mTSS data for Subjects who were Initially Randomized to Placebo and Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>15</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>16</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>(First Dose Date of Filgotinib + 14) - Study Day 1 + 1</td>
</tr>
<tr>
<td>Week 52*</td>
<td>On-Filgotinib Study Day</td>
<td>197</td>
<td>2</td>
</tr>
</tbody>
</table>

* For subjects who were on placebo and re-randomized their analysis window after Week 24 will be assigned according to On-Filgotinib Study Day, On-Filgotinib Lower Limit and On-Filgotinib Upper Limit.

The analysis windows for On-Treatment Healthcare Resource Utilization Questionnaire (HRUQ), Treatment Satisfaction Questionnaire for Medication (TSQM) and ECG data are provided in Table 3-7 and Table 3-8.
### Table 3-7. Analysis Visit Windows for On-Treatment HRUQ, TSQM and ECG Data for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>2</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>211</td>
</tr>
<tr>
<td>Week 36</td>
<td>253</td>
<td>212</td>
<td>309</td>
</tr>
<tr>
<td>Week 52</td>
<td>365</td>
<td>310</td>
<td>365</td>
</tr>
</tbody>
</table>

### Table 3-8. Analysis Visit Windows for On-Treatment HRUQ, TSQM and ECG Data for Subjects who were Initially Randomized to Placebo and Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>2</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>First Dose Date of Filgotinib - Study Day 1 + 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>On-Filgotinib Study Day</th>
<th>On-Filgotinib Lower Limit</th>
<th>On-Filgotinib Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 36*</td>
<td>85</td>
<td>2</td>
<td>141</td>
</tr>
<tr>
<td>Week 52*</td>
<td>197</td>
<td>142</td>
<td>197</td>
</tr>
</tbody>
</table>

* For subjects who were on placebo and re-randomized their analysis window after Week 24 will be assigned according to On-Filgotinib Study Day, On-Filgotinib Lower Limit and On-Filgotinib Upper Limit.

The analysis windows for on-treatment SF-36, FACIT-Fatigue, EQ-5D and WPAI-RA are provided in Table 3-9 and Table 3-10.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Table 3-9. Analysis Visit Windows for On-Treatment SF-36, FACIT-Fatigue, EQ-5D, and WPAI-RA for Subjects who were Initially Randomized to Filgotinib or Adalimumab, and Subjects who were Initially Randomized to Placebo and not Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 4</td>
<td>29</td>
<td>2</td>
<td>57</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>58</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>211</td>
</tr>
<tr>
<td>Week 36</td>
<td>253</td>
<td>212</td>
<td>309</td>
</tr>
<tr>
<td>Week 52</td>
<td>365</td>
<td>310</td>
<td>h 365</td>
</tr>
</tbody>
</table>

Table 3-10. Analysis Visit Windows for On-Treatment SF-36, FACIT-Fatigue, EQ-5D, and WPAI-RA for Subjects who were Initially Randomized to Placebo and Re-randomized

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 4</td>
<td>29</td>
<td>2</td>
<td>57</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>58</td>
<td>127</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>128</td>
<td>First Dose Date of Filgotinib - Study Day 1 + 1</td>
</tr>
</tbody>
</table>

* For subjects who were on placebo and re-randomized their analysis window after Week 24 will be assigned according to On-Filgotinib Study Day, On-Filgotinib Lower Limit and On-Filgotinib Upper Limit.

Vital signs, weight, ECG, lipids and safety laboratory data collected in the post-treatment follow-up period will also be summarized. The analysis window for the post-treatment follow-up period is defined as from (the last dose date of any study drug + 8 days) to (the last dose date of any study drug + 30 days). If multiple valid, nonmissing measurements exist in the post-treatment follow-up visit window, then the latest record will be selected for analysis. If the chronological order cannot be determined (eg, more than 1 record on the same day with time missing), for any given subject, the value with the worst severity will be selected for categorical variable, and the average will be taken for continuous variable. Data obtained after last dose date plus 30 days will be excluded from the summaries, but will be included in the listings.

For efficacy endpoints, the same analysis visit windows will be applied to all available data collected during standard of care.
3.8.3. Selection of Non-Efficacy Data in the Event of Multiple Records in an Analysis Visit Window

Depending on the statistical analysis method, single values may be required for each analysis window. For example, change from baseline by visit usually requires a single value, whereas a time-to-event analysis would not require 1 value per analysis window.

If multiple valid, nonmissing, continuous measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- In general, the baseline value will be the last nonmissing value on or prior to the first dose date of study drug, unless otherwise specified. If multiple measurements occur on the same day, the last nonmissing value prior to the time of first dose of study drug will be considered as the baseline value. If these multiple measurements occur at the same time or the time is not available, the average of these measurements (for continuous data) will be considered the baseline value.

- For postbaseline visits:
  - The record closest to the nominal day for that visit will be selected.
  - If there are 2 records that are equidistant from the nominal day, the later record will be selected.
  - If there is more than 1 record on the selected day, the average will be taken, unless otherwise specified.

If multiple valid, nonmissing, categorical measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- For baseline, the last available record on or prior to the date of the first dose of study drug will be selected. If there are multiple records with the same time or no time recorded on the same day, the value with the lowest severity will be selected (e.g., normal will be selected over abnormal for safety ECG findings).

- For postbaseline visits:
  - The record closest to the nominal day for that visit will be selected.
  - If there are 2 records that are equidistant from the nominal day, the later record will be selected.
  - If there is more than 1 record on the selected day, the value with the worst severity will be selected (e.g., abnormal will be selected over normal for safety ECG findings), unless otherwise specified.

The rules for selecting efficacy data in the event of multiple records in an analysis visit window are specified in Section 6.1.
4. **SUBJECT DISPOSITION**

4.1. **Subject Enrollment and Disposition**

A summary of subject enrollment will be provided by treatment group for each country within each geographic region and investigator within a country. The summary will present the number and percentage of subjects enrolled. For each column, the denominator for the percentage calculation will be the total number of subjects analyzed for that column.

A similar enrollment table will be provided by stratification factor stratum. The denominator for the percentage of subjects in the stratum will be the total number of enrolled subjects. If there are discrepancies in the value used for stratification assignment between the IWRS and the clinical database, the value collected in the clinical database will be used for the summary. A listing of subjects with discrepancies in the value used for stratification assignment between the IWRS and the clinical database at the time of data finalization will be provided.

The randomization schedule used for the study will be provided in a listing and as an appendix to the CSR.

A summary of subject disposition will be provided by treatment group. This summary will present the number of subjects screened, the number of subjects not randomized, the number of subjects who met all eligibility criteria but were not randomized with reasons subjects were not randomized, the number of subjects randomized, and the number of subjects in each of the categories listed below:

- Safety Analysis Set
- Full Analysis Set
- Per-Protocol Analysis Set
- PK Analysis Set
- PK Substudy Analysis Set
- Continuing study drug (Week 24 Analysis only)
- Completed study drug
- Continuing study (Week 24 Analysis only)
- Did not complete study drug with reasons for premature discontinuation of study drug
- Completed study
- Did not complete study with reasons for premature discontinuation from the study
Filgotinib GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

- Completed Week 24 study drug
- Prematurely discontinued study drug prior to Week 24
- Re-randomized

For the status of study drug and study completion and reasons for premature discontinuation, the number and percentage of subjects in each category will be provided. The denominator for the percentage calculation will be the total number of subjects in the Safety Analysis Set corresponding to that column.

For subjects who were on placebo, they will be considered as completed Week 24 study drug if they were re-randomized. For subjects who were on placebo but were not re-randomized, and subjects who remained on the same treatment, they will be considered as completed Week 24 study drug if (last dose date - Study Day 1 + 1) ≤ 156. Otherwise subjects will be considered as not completed Week 24 study drug.

The following by-subject listings will be provided by subject ID number in ascending order to support the above summary tables:

- Reasons for premature study drug or study discontinuation
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and kit ID of assigned study drugs

4.2. Extent of Drug Exposure and Adherence

Extent of exposure to study drug will be examined by assessing the total duration of exposure to study drug and the level of adherence to the study drug specified in the protocol.

4.2.1. Duration of Exposure to Study Drug

Total duration of exposure to any study drug will be defined as (last dose date of any study drug - first dose date of any study drug + 1), regardless of any temporary interruptions in study drug administration and will be expressed in weeks using up to 1 decimal place (e.g., 4.5 weeks).

For subjects with a partial last dosing date (i.e., month and year of last dose are known), the latest of the dispensing dates of study drug bottles, study drug start dates and end dates, and the imputed last dose date [day imputed as 15] will be used as the final imputed last dose date. If the subject died and the death date is complete (i.e., not partial date) and before the imputed last dose date, the complete death date will be used as the imputed last dose date.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

If only year is recorded (i.e., month and day of last dose are missing), the latest of the dispensing month of study drug bottles, study drug start month, and study drug end month will be used to impute the unknown last dose month. If the subject died and the death date has month and year available and before the imputed last dose month, then the month of death will be used instead. With the month imputed, the aforementioned method will be used to impute the last dose date.

If subjects are continuing on study drug at Week 24 analysis, the earliest of the date of death or data cutoff date for analysis will be used to impute the last dosing date for the calculation of duration of exposure to study drug.

The total duration of exposure to any study drug will be summarized using descriptive statistics and using the number (i.e., cumulative counts) and percentage of subjects exposed through the following time periods: Baseline (Day 1), Week 2 (Day 15), Week 4 (Day 29), Week 8 (Day 57), Week 12 (Day 85), Week 14 (Day 99), Week 16 (Day 113), Week 20 (Day 141), Week 24 (Day 169), Week 26 (Day 183), Week 30 (Day 211), Week 36 (Day 253), Week 44 (Day 309) and Week 52 (Day 365). A 7-day window is applied to the last planned on-treatment visit to match the protocol-specified visit window.

Summaries will be provided by treatment group for the Safety Analysis Set. No formal statistical testing is planned.

4.2.2. Adherence to Study Drug

Adherence will be calculated separately for filgotinib 200 mg/PTM (tablets), filgotinib 100 mg/PTM (tablets) and adalimumab/PTM (syringes).

The total number of tablets or syringes administered will be summarized using descriptive statistics.

The presumed total number of tablets or syringes administered to a subject will be determined by the data collected on the drug accountability CRF using the following formula:

\[
\text{Total Number of Tablets Administered} = \sqrt{\text{No. of Tablets Dispensed}} - \sqrt{\text{No. of Tablets Returned}}
\]

\[
\text{Total Number of Syringes Administered} = \sqrt{\text{No. of Syringes Dispensed}} - \sqrt{\text{No. of Syringes Returned}}
\]

If a bottle or a packaging kit is dispensed and the bottle or packaging kit is returned empty, then the number of tablets/syringes returned will be entered as zero. If a bottle or a packaging kit is dispensed but not returned (missing), then the number of tablets/syringes returned will be counted as zero.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

Version 1.0

4.2.2.1. On-Treatment Adherence

The level of on-treatment adherence to the study drug will be determined by the total amount of study drug administered relative to the total amount of study drug expected to be administered during a subject’s actual on-treatment period based on the study drug regimen.

The level of on-treatment adherence will be expressed as a percentage using the following formula:

\[
\text{On-Treatment Adherence (\%) = } \left( \frac{\text{Total Amount of Study Drug Administered}}{\text{Study Drug Expected to be Administered on Treatment}} \right) \times 100
\]

Note: If calculated adherence is greater than 100%, the result will be set to 100%.

Study drug expected to be administered for filgotinib 200 mg/PTM (tablets) = 1B total duration of exposure to any study drug (days).

Study drug expected to be administered for filgotinib 100 mg/PTM (tablets) = 1B total duration of exposure to any study drug (days).

Study drug expected to be administered for adalimumab/PTM (syringes) = ceil[total duration of exposure to any study drug (days) /14].

For subjects in placebo group who were re-randomized, the switched study drug expected to be administered is defined as:

Study drug expected to be administered for PTM (tablets) = 1B ((first dose date of filgotinib - 1 - first dose date of any study drug) + 1).

Study drug expected to be administered for filgotinib 100 mg/ filgotinib 200 mg = 1B ((last dose date of any study drug - first dose date of filgotinib) + 1).

On-treatment adherence will be calculated up to Week 12, up to Week 24 and up to Week 52, respectively.

Descriptive statistics for the level of on-treatment adherence with the number and percentage of subjects belonging to adherence categories (eg, < 80%, h 80 to < 90%, h 90%) will be provided by treatment group for the Safety Analysis Set.

Categorical displays will be presented for the number of subjects who are at least 80% adherent to their study drug regimen (ie, adherence is h 80% for each study drug).

No formal statistical testing is planned.
A by-subject listing of study drug administration and drug accountability will be provided separately by subject ID number (in ascending order) and visit (in chronological order).

4.3. Protocol Deviations

Subjects who did not meet the eligibility criteria for study entry, but enrolled in the study will be summarized regardless of whether they were exempted by the sponsor or not. The summary will present the number and percentage of subjects who did not meet at least 1 eligibility criterion and the number of subjects who did not meet specific criteria by treatment group based on the All Randomized Analysis Set. A by-subject listing will be provided for those subjects who did not meet at least 1 eligibility (inclusion or exclusion) criterion. The listing will present the eligibility criterion (or criteria if more than 1 deviation) that subjects did not meet and related comments, if collected.

Protocol deviations occurring after subjects entered the study are documented during routine monitoring. The number and percentage of subjects with important protocol deviations by deviation reason (eg, nonadherence to study drug, violation of select inclusion/exclusion criteria) will be summarized by treatment group for the All Randomized Analysis Set. A by-subject listing will be provided for those subjects with important protocol deviation.
5. BASELINE CHARACTERISTICS

5.1. Demographics and Other Baseline Characteristics

Subject demographic and other baseline characteristics variables will be summarized by treatment group and overall using descriptive statistics for continuous variables, and using number and percentage of subjects for categorical variables. The summary of demographic and other baseline characteristics data will be provided for the Safety Analysis Set for the following:

- Age (on the first dose date of any study drug)
- Age group (< 65 years, ≥ 65 years)
- Sex at birth (male, female)
- Race
- Ethnicity (Hispanic or Latino, not Hispanic or Latino)
- Geographic region and country
- Weight (kg)
- Height (cm)
- Body mass index (BMI; in kg/m²)
- Smoking status

A by-subject demographic and other baseline characteristics listing, including the informed consent date, will be provided by subject ID number in ascending order.

5.2. Baseline Disease Characteristics

Baseline disease characteristics include:

- Duration of RA from diagnosis (years)
  Calculated as \((\text{first dose date} - \text{date of initial diagnosis}) + 1 \text{ day}) / 365.25\). If the date of initial diagnosis is incomplete, then the following rules will be applied:
  - missing day: use the first of the month
  - missing month: use January
- Presence of RF (Yes/No)
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

- Presence of anti-CCP Ab (Yes/No)
- Presence of both RF and anti-CCP Ab (Yes/No)
- Prior exposure to bDMARDs (Yes/No)
- Concurrent oral corticosteroids use on the first dosing date (Yes/No): n (%)
  - Oral corticosteroids dose, mg/day, expressed as prednisone-equivalent dose
- Concurrent MTX use on the first dosing date (Yes/No): n (%)
  - Dose of MTX, mg/week
- Concurrent anti-malarials use on the first dosing date (Yes/No)
- Swollen joint count based on 66 joints (SJ C66)
- Tender joint count based on 68 joints (TJC68)
- Swollen joint count based on 28 joints (SJ C28)
- Tender joint count based on 28 joints (TJC28)
- HAQ-DI total score
- DAS28(CRP)
- SF-36 physical component summary (PCS) score
- SF-36 mental component summary (MCS) score
- FACIT-Fatigue
- Subject’s pain assessment (by visual analog scale [VAS] in mm)
- Subject’s global assessment of disease activity [SGA] (by VAS in mm)
- Physician’s global assessment [PGA] (by VAS in mm)
- Simplified Disease Activity Index (SDAI)
- Clinical Disease Activity Index (CDAI)
- hsCRP (mg/L)
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

- hsCRP ≥ 6 mg/L (Y es/No)
- mTSS
  - Erosion score
  - Joint space narrowing (JSN) score
- Subject with erosion defined as erosion score > 0 (Y es/No)

A list of anti-malarials is provided in Appendix 1. The baseline disease characteristics will be summarized by treatment group and overall using descriptive statistics for continuous variables and using number and percentage of subjects for categorical variables. The summary of baseline disease characteristics will be provided for the Safety Analysis Set.

A by-subject listing of baseline disease characteristics will be provided by subject ID number in ascending order.

5.3. Medical History

Medical history collected at screening will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) 21.0.

Medical history will be summarized by system organ class (SOC), preferred term (PT), treatment group, and overall. Subjects who report 2 or more medical history items that are coded to the same SOC and/or PT will be counted only once by the unique coded term in the summary.

In addition, numbers and percentages of subjects who have any first degree relatives that had experienced myocardial infarction or stroke before the age of 50 years, experienced myocardial infarction or stroke before the age of 50 years, or have diabetes (type I or II) will be summarized.

The summary will be provided for the Safety Analysis Set. No formal statistical testing is planned.

A by-subject listing of medical history will be provided by subject ID number.
6. **EFFICACY ANALYSES**

6.1. **General Considerations**

The primary analysis set for efficacy analyses will be the FAS, defined in Section 3.1.2.

Efficacy analysis will generally be conducted on the following 2 datasets:

1) On-treatment data as specified in Section 3.8.2 with exclusion of the observations described as the following. This is the primary analysis dataset for efficacy analysis.

   - For subjects who continued study drug BUT did not achieve at least 20% improvement from Baseline in both SJ C66 and TJ C68 at
     - Nominal Week 14 visit, then efficacy data collected after nominal Week 14 visit will be excluded from the primary analysis, OR
     - Nominal Week 12 visit if subjects missed Week 14 visit or the response status cannot be determined due to missing SJ C66 or TJ C68 value at Week 14, then efficacy data collected after nominal Week 16 visit will be excluded from the primary analysis, OR
     - Nominal Week 16 visit if subjects missed both Week 12 and Week 14 visits or the response status cannot be determined due to missing SJ C66 or TJ C68 value at both Week 12 and Week 14 visits, then efficacy data collected after nominal Week 16 visit will be excluded from the primary analysis

   If subjects missed all three visits (Week 12, Week 14 and Week 16) or the response status cannot be determined at all of these three visits, all the efficacy data collected will be included for analysis

   - Subjects who continued on study drug will be evaluated for loss of therapeutic response from nominal Week 30 visit through nominal Week 52 visit. For subjects who failed to maintain at least a 20% improvement from Baseline in SJ C66 and TJ C68 (which is confirmed at 2 consecutive visits, and a missed visit will not count against them), efficacy data collected after the second consecutive visit will be excluded from the primary analyses.

2) All available data (including data collected under standard of care).

**Estimands:**

Two efficacy estimands, composite estimand and treatment policy estimand are defined for the primary and key secondary efficacy endpoints, respectively.
The **Composite Estimand** is defined as following. This is the primary estimand for efficacy endpoints.

1) Population: Subjects in the FAS.

2) Variable: Primary and key secondary endpoints as well as some other secondary endpoints (see details in Section 6.2, 6.3 and 6.4)

3) Intercurrent events: The following intercurrent events are taken into account:
   - If a subject takes standard of care medications due to inadequate response or discontinues from study treatment, treat data as missing from the point of inadequate response or study drug discontinuation (i.e., on-treatment data defined above).
   - If a subject discontinues from the study, treat data as missing from the point of loss to follow-up onward.

4) Population-level summary: Proportion difference for binary efficacy endpoints or mean difference in change from baseline for continuous efficacy endpoints between each filgotinib group and placebo group or adalimumab group if applicable.

The **Treatment-Policy Estimand** is defined as following. This is the secondary estimand for efficacy endpoints.

1) Population: Subjects in the FAS.

2) Variable: Primary and key secondary endpoints as well as some other secondary endpoints (see details in Section 6.2, 6.3 and 6.4)

3) Intercurrent events: The occurrence of an intercurrent event is irrelevant. All observed values will be used regardless of occurrence of an intercurrent event.
   - All observed values will be used regardless of occurrence of an intercurrent event (i.e., all available data defined above).
   - If a subject discontinues from the study, treat data as missing from the point of loss to follow-up onward.

4) Population-level summary: Proportion difference for binary efficacy endpoint or mean difference in change from baseline for continuous efficacy endpoints between each filgotinib group and placebo group or adalimumab group if applicable.

The other secondary endpoints specified in both estimands include ACR50, ACR70, and the individual components of the ACR response.
To handle the missing data caused by the intercurrent events, the following analyses will be performed for both estimands:

- For binary endpoints, missing data will be treated as non-responder for the primary analysis. Sensitivity analyses including observed data only, last observation carry forward (LOCF), multiple imputation, and tipping point methods will be applied.

- For continuous endpoints, missing data will be handled using a mixed-effects model for the primary analysis. Sensitivity analyses, including using multiple imputation and tipping point methods, will be applied.

In addition, analyses will be performed for the primary and key secondary endpoints by replacing the population in both estimands with the Per-Protocol analysis set as defined in Section 3.1.3. For these analyses, the binary missing data will be imputed as non-responder and the continuous missing data will be handled using a mixed-effects model.

**Selection of Efficacy Data in the Event of Multiple Records in an Analysis Visit Window**

If multiple valid, nonmissing efficacy measurements exist in an analysis window, records will be chosen based on the following rules if a single value is needed:

- The record closest to the nominal day for that visit will be selected

- If there are 2 records that are equidistant from the nominal day, or more than 1 record (with time known) on the selected day, the latest record will be taken

- If chronological order cannot be determined (eg, more than 1 record on the same day with time missing), for any given subject, the worst outcome will be selected.

**Calculation of Composite Endpoints**

For the calculation of composite endpoints including DAS28(CRP), ACR20/50/70, ACR-N, SDAI, and CDAI, we use the following steps unless otherwise specified:

- Step 1: Assign individual components to analysis visit windows defined in Section 3.8.2

- Step 2: Within each analysis visit window, select the component-level data based on the rules for selecting efficacy data as above

- Step 3: Calculate the composite endpoint based on the selected component-level data in Step 2.
Missing Data Imputation

Below are the descriptions for the imputation methods that will be used throughout the efficacy analyses:

- **Observed case (OC):** Missing values remain missing. For the categorical composite endpoints, in the case that some components are missing, the composite endpoint assessment will be derived based on the non-missing components. If non-missing components are not sufficient to determine final composite endpoint, then the composite endpoint will be set as missing. For continuous composite endpoints (including ACR-N), if any components are missing, the composite endpoints will be set as missing.

- **Last observation carried forward (LOCF):** Baseline measurements will not be carried forward to postbaseline. Only postbaseline measurements will be LOCF. For the composite endpoints, the last nonmissing postbaseline observation will be carried forward to subsequent visits for each individual component first, and then calculate the composite endpoints using individual components imputed by LOCF as described above. If a subject does not have a nonmissing observed record for a postbaseline visit, the last postbaseline record prior to the missed visit will be used for this postbaseline visit. If the last nonmissing observation prior to the missing visits cannot be determined due to multiple measurements occurring at the same time or the time not available within the same day, the worst outcome will be used for LOCF. If missing components still exist after LOCF, the composite endpoints will be calculated using the same rules as described in OC.

- **Non-responder imputation (NRI):** For all binary response measurements, starting from OC, all missings will be set as non-responders.

If a subject only has baseline measurements, LOCF and OC analyses will not include this subject. But this subject will be treated as non-responder in NRI analyses.

6.1.1. Tender/Swollen Joint Counts (TJC68/SJC66)

Tender joint count based on 68 joints (TJC68) and swollen joint count based on 66 joints (SJC66) will be collected during the course of the study. The assessment for each joint will be from the following selections: Tender Only, Swollen Only, Tender and Swollen, Joint Non-Evaluable or Missing, or Not Tender or Swollen.

Individual joint with missing assessment will not be imputed. If at least half of the joints are assessed at a given visit, the prorated tender and swollen joint counts will be calculated using the following formula:

\[ TJC68 = \frac{\text{Total number of tender joints}}{68 - (\text{Number of nonevaluable or missing joints out of 68 joints})} \times 68 \]

\[ SJC66 = \frac{\text{Total number of swollen joints}}{66 - (\text{Number of nonevaluable or missing joints out of 66 joints})} \times 66 \]
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

If less than half of joints are assessed at a given visit, joint counts are treated as missing for that visit.

A more abbreviated assessment considering 28 joints as listed in Table 6-1 for both tenderness and swelling will also be conducted (as part of the TJC68 and SJC66 assessment), denoted as TJC28 and SJC28, respectively.

**Table 6-1. Composition of the 28 Joints**

<table>
<thead>
<tr>
<th>Joints</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder Joints (Left and Right)</td>
<td>2</td>
</tr>
<tr>
<td>Elbow Joints (Left and Right)</td>
<td>2</td>
</tr>
<tr>
<td>Wrist Joints (Left and Right)</td>
<td>2</td>
</tr>
<tr>
<td>Metacarpophalangeal Joints I-V (Left and Right) * hands only</td>
<td>10</td>
</tr>
<tr>
<td>Proximal Interphalangeal Joints I-V (Left and Right) * hands only</td>
<td>10</td>
</tr>
<tr>
<td>Knee Joints (Left and Right)</td>
<td>2</td>
</tr>
</tbody>
</table>

If there exist non-evaluable or missing joints among the 28 joints, similar prorated tender and swollen joint counts will be calculated as follows:

\[
TJC28 = \frac{\text{Total number of tender joints}}{28 - \text{(Number of nonevaluable or missing joints out of 28 joints)}} \times 28
\]

\[
SJC28 = \frac{\text{Total number of swollen joints}}{28 - \text{(Number of nonevaluable or missing joints out of 28 joints)}} \times 28
\]

If less than half of the 28 joints are assessed at a given visit, TJC28 and SJC28 are treated as missing for that visit.

**6.1.2. Global Assessment of Disease Activity**

Subject’s Global Assessment of Disease Activity (SGA) and Physician’s Global Assessment of Disease Activity (PGA) based on a 0-100 mm visual analog scale (VAS) will be recorded during the study, with 0 indicating "no disease activity" and 100 indicating "maximum disease activity" (or similar description of disease severity).
6.1.3. Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), administered by the subject. Responses in each functional category are collected as: without any difficulty; with some difficulty; with much difficulty; unable to do a task in that area and with or without aids or devices. The HAQ-DI score ranges from 0 (no disability) to 3 (completely disabled), when 6 or more categories are non-missing. Detailed algorithm for calculating HAQ-DI score is described in Appendix 2.

HAQ-DI also includes a separate pain assessment and subject will be requested to mark the severity of the pain in the past week on a 0-100 mm VAS, with 0 indicating `no pain`, and 100 indicating `severe pain`.

6.2. Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of subjects who achieve ACR20 response at Week 12.

6.2.1. Definition of the Primary Efficacy Endpoint

A subject achieves ACR20 response when this subject has

- h 20% improvement from baseline in TJC68, AND
- h 20% improvement from baseline in SJC66, AND
- h 20% improvement from baseline in at least 3 of the following 5 items:
  1) PGA
  2) SGA
  3) Subject’s pain assessment
  4) HAQ-DI score
  5) hsCRP

Percent improvement from baseline at a postbaseline visit is calculated as follows for all 7 components mentioned above:

\[
\text{% improvement} = \left(\frac{\text{baseline value} - \text{postbaseline value}}{\text{baseline value}}\right) \times 100
\]

If the baseline value is 0 then the percent improvement from baseline is set to missing.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

In the case that some ACR20 components are missing, the ACR20 assessment will be based on the non-missing components. If non-missing components are not sufficient to determine ACR20 response, then the ACR20 response will be considered as missing.

6.2.2. Statistical Hypothesis for the Primary Efficacy Endpoint

In the primary analysis, the ACR20 response rate at Week 12 in the filgotinib 200 mg group will be compared to placebo group for a superiority test at the 2-sided 0.05-level. If we denote the ACR20 response rate at Week 12 in the filgotinib 200 mg group and placebo group as $P_1$ and $P_2$ respectively, the null and alternative hypotheses for the superiority test on the primary efficacy endpoint are as follows:

$H_0: P_1 = P_2$

vs

$H_1: P_1 > P_2$

6.2.3. Primary Analyses of the Primary Efficacy Endpoint

To test for superiority of filgotinib 200 mg group versus placebo control group in proportion of subjects who achieve ACR20 at Week 12, a logistic regression analysis with treatment groups and stratification factors in the model will be used. The model will include all treatment groups. Subjects who do not have sufficient measurements to establish efficacy at Week 12 will be considered as failures (i.e., NRI). The $p$-value from the logistic regression model for testing the superiority of filgotinib 200 mg as compared to placebo will be provided. The 2-sided 95% confidence interval (CI) of the ACR20 response rate at Week 12 based on normal approximation with the sample variance will be provided for each treatment group. In addition, non-stratified ACR20 response rate difference along with its 95% CI calculated based on the normal approximation with the pooled sample variance will be provided. Appendix 8 provides samples SAS model statements for constructing the confidence interval for the proportion.

The comparisons of filgotinib 100 mg groups versus placebo group in the proportion of subjects achieving ACR20 at Week 12 will also be conducted with the similar logistic regression model with NRI as aforementioned.

The number and percentage of ACR20 non-responders observed and non-responders due to missing at Week 12 will also be summarized respectively by treatment.

For subjects with observed ACR20 outcomes (i.e, responders and observed non-responders) at Week 12, actual values and change from Baseline at Week 12 in individual components, including TJ C68, SJ C66, SGA, PGA, Subject's pain assessment, hsCRP, will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment.
The proportions of subjects achieving ACR20 (using NRI) will be plotted over time from Baseline through Week 24 by treatment and by visit.

### 6.2.4. Sensitivity Analysis of the Primary Endpoint

Sensitivity analyses of the primary efficacy endpoint will be performed.

#### 6.2.4.1. Per-protocol Analyses

To evaluate the impact of study conduct on the primary analysis, the proportion of subjects who achieve ACR20 at Week 12 will be analyzed based on the PP Analysis Set as defined in Section 3.1.3 using the logistic regression analysis with NRI for the comparisons of each filgotinib group versus the placebo group. The p-value from the logistic regression model will be provided. The 2-sided 95% CI of the ACR20 response rate at Week 12 based on normal approximation with the sample variance for each treatment group will be provided. In addition, the non-stratified ACR20 response rate difference along with its 95% CI calculated based on the normal approximation with the pooled sample variance will also be provided.

#### 6.2.4.2. Missing Data Imputation Analyses

To evaluate the impact of missing data on the ACR20 response rate at Week 12, the following missing value imputation methods will be used:

- The analysis specified in Section 6.2.3 will be performed by using OC and LOCF methods as described in Section 6.1.

- Multiple imputation (MI): The MI procedure replaces each missing binary ACR20 value with a set of plausible values that represent the uncertainty about the right value to impute. Twenty imputed datasets will be generated based on logistic regression models with starting seed 12345. These multiple imputed data sets are then analyzed by using the same method for the primary analysis for complete data as specified in Section 6.2.3. The results from each set of imputed data sets will then be combined using Rubin’s rule (Rubin 1987). The stratification factors will be included in the imputation model as covariates and data at post-baseline visits up to the time point of interest will be included in the longitudinal model.
Filgotinib GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis Version 1.0

- Tipping point analysis: To assess the robustness of analysis results under MNAR (missing not at random) assumption, a delta-adjusting pattern-mixture approach for tipping point analysis [Ratitch 2013] will be conducted for the primary efficacy endpoint. The impact from missing data on the comparisons in proportion of subjects achieving ACR20 at Week 12 between filgotinib groups and placebo control group will be examined, by adjusting for stratification factors. The proposed method will perform a series of analyses with a range of different values of the shift parameter applied to the imputed datasets at which the conclusion about the statistical significance of the estimated treatment effect will be altered. Specifically, a tipping point analysis is characterized by a pair of sequences (one sequence associated with missing data in the placebo group and the other associated with missing data in the filgotinib group). In the first step for both sequences, the missing data in both treatment groups are imputed under the MAR assumption. For each subsequent step of the placebo sequence, the missing data is imputed assuming the missing data is incrementally more favorable than in the previous step of that sequence. Similarly, for each subsequent step of the active sequence, the missing data is imputed assuming the missing data is incrementally less favorable than in the previous step of that sequence. Appendix 6 provides sample SAS code statements for the tipping point analysis with starting seed. Each value is classified as either: altering the study’s conclusion or tips: keeping the study’s conclusion unchanged. The tipping points that alter the statistical conclusion will be provided. For each value, 20 imputed data sets will be generated. The same analysis method for the primary analysis as specified in Section 6.2.3 will be applied when analyzing adjusted data generated under the different values.

6.2.5. Subgroup Analysis of the Primary Endpoint

Subgroup analyses comparing each filgotinib dose group to the placebo group will be performed at Week 12 and Week 24 for the primary endpoint, for the subgroups specified in Section 3.4.

The proportion of subjects who achieve ACR20 will be analyzed using the Fisher’s exact test based on the NR1 method for treatment group comparison. The number and percentage of subjects with ACR20 will also be provided for each treatment group within the subgroups.

6.3. Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints are:

- Change from Baseline in HAQ-DI score at Week 12
- The proportion of subjects who achieve DAS28(CRP) < 2.6 at Week 12
- Change from Baseline in mTSS at Week 24
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis  Version 1.0

- The proportion of subjects who achieve DAS28(CRP) ≥ 3.2 at Week 12
- Change from Baseline in SF-36 physical component summary (PCS) score at Week 12
- Change from Baseline in FACIT-Fatigue score at Week 12

6.3.1. Definition of Key Secondary Efficacy Endpoints

6.3.1.1. Disease Activity Score for 28 Joint Count using CRP

The DAS28(CRP) score is calculated as follows:

\[
DAS28(CRP) = 0.56 \sqrt{TJC28} + 0.28 \sqrt{SJC28} + 0.36 \ln(CRP + 1) + 0.014 \times SGA + 0.96,
\]

where

- CRP = hsCRP measurement (mg/L)
- SGA = subject’s global assessment of disease activity on a 0-100 VAS

Higher DAS28(CRP) value indicates more severe disease activity.

No component-level imputation will be performed for the calculation of DAS28(CRP) for the primary analyses. If any components are missing, the DAS28(CRP) will be set as missing.

6.3.1.2. Modified Total Sharp Score (mTSS)

Subject’s radiographs of bilateral hands, wrists and feet will be taken at screening and protocol specified visits. The radiographs will be evaluated through central review by independent joint assessors using the modified total Sharp Score (mTSS) method. The mTSS (range [0, 448]) is defined as the erosion score (range [0, 280]) plus the joint space narrowing (JSN) score (range [0, 168]). Detailed algorithm for calculating mTSS, erosion score and JSN score is described in Appendix 3.

6.3.1.3. 36-Item Short-form Health Survey

The SF-36 Version 2 is a 36-item, self-reported, generic, comprehensive, and health-related quality of life questionnaire that yields an 8 health domains (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health). Each domain is scored by summing the individual items and transforming the scores into a 0 to 100 scale with higher scores indicating better health status or functioning. In addition, 2 summary scores, the physical component summary (PCS) score and the mental component summary (MCS) score will be evaluated based on the 8 SF-36 domains.
6.3.1.4. FACIT-Fatigue

The FACIT-Fatigue scale is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning in the past 7 days. The FACIT-Fatigue uses 0 (‘not at all’) to 4 (‘very much’) numeric rating scales. Negatively stated items are reversed by subtracting the response from ‘4’ before being added to obtain a total score. Scores range from 0 to 52 with higher scores indicating less fatigue. In the case of missing response for some items in the questionnaire, if at least half of the items (ie, 7 of 13 items) were answered at a given visit, the prorated score will be calculated and used in the analysis.

6.3.2. Analysis Methods for Key Secondary Efficacy Endpoints

Hypothesis testing on the key secondary efficacy endpoints will commence after the testing on the primary efficacy endpoint reaches statistical significance, and will be performed according to the hierarchical testing principle at the 2-sided 0.05 level as described in Section 3.5. If a null hypothesis is not rejected, formal sequential testing will be stopped and only nominal significance will be reported for the remaining hypotheses.

For non-inferiority test of the proportion of subjects who achieve DAS28(CRP) ≥ 3.2 at Week 12, the approach proposed by (Liu 2014) will be used to demonstrate that each filgotinib dose preserves more than 50% of the effect of adalimumab on the response rate of DAS28 (CRP) ≥ 3.2 at Week 12 using NRI. Let $\pi_T$, $\pi_C$ and $\pi_P$ denote the true response rates of DAS28 (CRP) ≥ 3.2 for filgotinib, adalimumab, and placebo respectively, at Week 12, with corresponding variances $\sigma_T^2$, $\sigma_C^2$ and $\sigma_P^2$. The non-inferiority null hypothesis is

$$H_{0 NI}: \frac{\pi_T - \pi_P}{\pi_C - \pi_P} \geq 0.5$$

According to (Liu 2014), non-inferiority (filgotinib preserves more than 50% of the effect of adalimumab) at 2-sided 0.05 level will be claimed if $Z_{NI} > z_*$, where

$$Z_{NI} = \frac{\hat{\pi}_T - 0.5\hat{\pi}_C - (1-0.5)\hat{\pi}_P}{\sqrt{\hat{\sigma}_T^2 + 0.5\sigma_C^2 + (1-0.5)^2\hat{\sigma}_P^2}}$$

and $z_*$ is the 97.5 quantile of the standard normal distribution.

Note that the ‘hat’ (^) denotes the estimated values of each parameter at Week 12.

The non-inferiority test of the proportion of subjects who achieve DAS28(CRP) < 2.6 at Week 12 will be performed.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis  Version 1.0

For the superiority test of the proportion of subjects who achieve DAS28(CRP) < 2.6 at Week 12, the similar logistic regression analysis with the same model specification as the primary endpoint analysis described in Section 6.2.3 will be adopted for the treatment comparisons. The model will include all treatment groups. The NRI method will be used to impute missing value. Comparison will be made between each filgotinib dose group and placebo group. The p-value from the logistic regression model will be reported for statistical inference. The 2-sided 95% CI of response rate based on normal approximation with the sample variance will be provided for each treatment group. In addition, non-stratified response rate difference along with its 95% CI calculated based on the normal approximation with the pooled the sample variance will be provided. The superiority test of the proportion of subjects who achieve DAS28(CRP) \( \geq 3.2 \) or DAS28(CRP) < 2.6 at Week 12 will be performed in the similar way, and the comparison will be made between each filgotinib dose group and adalimumab group.

The change from Baseline in HAQ-DI at Week 12 will be analyzed using the mixed-effects model for repeated measures (MMRM) that includes data at postbaseline visits up to the time point of interest. Subjects that have a baseline value and at least 1 postbaseline value are included in the analysis. The MMRM models will be used to evaluate treatment effect on change score from Baseline, with baseline value, stratification factors, treatment, visit, and treatment by visit interaction included as fixed effects and subject being the random effect. The MMRM model will include all treatment groups. An unstructured variance-covariance matrix will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. Missing change scores due to missing study visits, early withdrawal or treatment reassignment will not be otherwise imputed using the MMRM approach. The least squares (LS) means and 95% CIs of the difference in mean change from Baseline in HAQ-DI between each filgotinib dose group and placebo group from MMRM will be provided.

The change from Baseline in mTSS, erosion score and JSN score at Week 24, and change from Baseline in SF-36 PCS score and FACIT-Fatigue score at Week 12 will be analyzed using similar MMRM method as HAQ-DI, by including data at postbaseline visits up to the time point of interest. To test for a treatment difference between each filgotinib dose group and placebo, LS mean difference along with 95% CI and p-values from the MMRM model will be presented.

The proportions of subjects achieving DAS28(CRP) \( \geq 3.2 \) (using NRI), and proportions of subjects achieving DAS28(CRP) < 2.6 (using NRI) will be plotted over time by treatment. Plots of mean \( \pm \) SD of changes from Baseline in HAQ-DI, mTSS, erosion score, JSN score, SF-36 PCS and FACIT-Fatigue over time will be presented respectively. Cumulative percentile plots will be provided for change from baseline in mTSS.

6.3.3. Sensitivity Analysis for Key Secondary Efficacy Endpoints

6.3.3.1. Per-Protocol Analysis

The analyses of the key secondary efficacy endpoints (except for the change from baseline in mTSS at Week 24) will be repeated using the PP Analysis Set specified in Section 3.1.3.
6.3.3.2. Missing Data Imputation Analyses

The following imputation methods will be explored:

- Change from Baseline in HAQ-DI score at Week 12
  1) MI: All subjects with baseline measurement will be included. The MI procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. Twenty imputed datasets will be generated based on linear regression models on observed HAQ-DI scores with starting seeds 123 and 465. These 20 imputed data sets are then analyzed by using the analysis method specified in Section 6.3.2 for complete data. The results from each imputed data set will then be combined using Rubin’s rule (Rubin 1987). Stratification factors and baseline values will be included in the imputation model as covariates, and data at baseline and all post-baseline visits up to the time point of interest visit will be included in the longitudinal model.

  2) Tipping point analysis: All subjects with baseline measurement will be included. To access the robustness of analysis results for change from Baseline in HAQ-DI at Week 12 under MMAR assumption, a set of the shift parameters that adjust the imputed values will be examined (Yuan 2014). The shift parameter that alters the study conclusion for the hypothesis testing for change from Baseline in HAQ-DI at Week 12 under the MAR assumption will be reported as the tipping point and provided. The tipping point analysis will be conducted by iteratively assigning plausible outcomes to missing values for subjects in different treatment group independently until the conclusion is reversed (eg, analyses are no longer statistically significant). Appendix 7 provides sample SAS model statements for the tipping point analysis with starting seeds. The same analysis method for the primary analysis as specified in Section 6.3.2 will be applied when analyzing adjusted data generated under each plausible shift parameter.

- Proportion of subjects who achieve DAS28(CRP) < 2.6 at Week 12
  1) The analysis using the same logistic regression model specified in Section 6.3.2 will be performed by using OC and LOCF methods as described in Section 6.1.

  2) MI: All subjects with baseline measurement will be included. The MI procedure replaces each missing composite DAS28(CRP) value with a set of plausible values that represent the uncertainty about the right value to impute. Baseline value and stratification factors will be included in the imputation model as covariates, and data at baseline and post-baseline visits up to the time point of interest will be included in the longitudinal model. The multiple imputed datasets will be generated. These multiple imputed data sets are then used to identify subjects with DAS28(CRP) < 2.6 at Week 12, and analyzed by using the analysis method specified in Section 6.3.2 for complete data. The results from each set of imputed data sets will then be combined using Rubin’s rule (Rubin 1987).

  3) Tipping point analysis: Similar modified tipping point method as described in Section 6.2.4.2 will be applied to DAS28(CRP) < 2.6 response rate at Week 12.
• Change from Baseline in mTSS at Week 24
  1) MI: Similar MI procedure used for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.
  2) Tipping point analysis: Similar tipping point analysis method used for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.

• Non-inferiority analysis of the proportion of subjects who achieve DAS28(CRP) \( \geq 3.2 \) or DAS28(CRP) < 2.6 at Week 12
  1) The non-inferiority analysis using the same test specified in Section 6.3.2 will be performed by using OC and LOCF methods as described in Section 6.1.
  2) MI: Similar MI procedure used for superiority test of the proportion of subjects who achieve DAS28(CRP) < 2.6 at Week 12 will be performed. The multiple imputed data sets will be analyzed by using the non-inferiority analysis specified in Section 6.3.2 for complete data. The results from each set of imputed data sets will then be combined.
  3) Tipping point analysis: Similar modified tipping point method as described in Section 6.2.4.2 will be applied to DAS28(CRP) \( \geq 3.2 \) or DAS28(CRP) < 2.6 response rate at Week 12. The non-inferiority analysis specified in Section 6.3.2 will be applied when analyzing adjusted data generated under the different values.

• Change from Baseline in SF-36 PCS at Week 12
  1) MI: Similar MI procedure used for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.
  2) Tipping point analysis: Similar tipping point analysis method as for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.

• Change from Baseline in FACIT-Fatigue at Week 12
  1) MI: Similar MI procedure used for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.
  2) Tipping point analysis: Similar tipping point analysis method as for analyzing the change from Baseline in HAQ-DI score at Week 12 will be performed.

6.3.3.3. Sensitivity Analyses of Non-inferiority Test
For non-inferiority test of the proportion of subjects who achieve DAS28(CRP) \( \geq 3.2 \) or DAS28(CRP) < 2.6 at Week 12, different margins that are higher than 0.5, and up to 0.99 will be used in the non-inferiority analysis. The null hypothesis is

\[
H_{0,NI} : \frac{\pi_T - \pi_P}{\pi_C - \pi_P} = \eta
\]
Filgotinib
GS-US-417-0301
Statistical Analysis Plan – Week 24 Analysis and Final Analysis Version 1.0

According to {Liu 2014}, non-inferiority with margin $δ$ at 2-sided 0.05 level will be claimed if $Z_{NI} > z_α$, where

$$Z_{NI} = \frac{\bar{Y}_p - \bar{Y}_c - (1 - \delta)\delta_p}{\sqrt{\delta_p^2 + (1 - \delta)^2\delta_c^2}}$$

and $z_α$ is the 97.5 quantile of the standard normal distribution.

6.3.4. **Subgroup Analyses for Key Secondary Efficacy Endpoints**

Subgroup analyses comparing each filgotinib dose group to the placebo group will be performed at Week 12 and Week 24 for the key secondary efficacy endpoints, for the subgroups specified in Section 3.4.

The subgroup analysis for the proportion of subjects who achieve DAS28(CRP) < 2.6 will be performed using the Fisher’s exact test based on the NRI method. The number and percentage of subjects with DAS28(CRP) < 2.6 will be provided for each treatment group within the subgroups.

The change from Baseline in HAQ-DI will be analyzed using the MMRM method that includes data at postbaseline visits up to the time point of interest with baseline value, treatment, visit, and treatment by visit as fixed effects and subjects being the random effect. The LS mean, LS mean difference, SE and 95% CI will be presented. The change from Baseline in mTSS, change from Baseline in SF-36 PCS and change from Baseline in FACIT-Fatigue will be analyzed similarly using the MMRM model, respectively. Descriptive statistics for actual values and change from baseline will also be presented for each treatment group within the subgroups.

6.4. **Other Secondary Efficacy Endpoints**

Other secondary efficacy endpoints include:

- The proportion of subjects who achieve ACR50 and ACR70 at Weeks 4, 12, 24 and 52, ACR20 at Weeks 4, 24, and 52, and ACR20/50/70 over time from Day 1 through Week 52
- Change from Baseline in individual components of the ACR response at Weeks 4, 12, 24, and 52 and over time from Day 1 through Week 52
- The proportion of subjects who achieve decrease in HAQ-DI of $\frac{1}{2}$ 0.22 at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in DAS28(CRP) at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28(CRP) $\frac{1}{2}$ 3.2 at Weeks 4, 24, and 52, and over time from Day 1 through Week 52
- The proportion of subjects who achieve DAS28(CRP) < 2.6 at Weeks 4, 24, and 52, and over time from Day 1 through Week 52
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

- ACR-N and EULAR response at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in Clinical Disease Activity Index (CDAI) at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52
- Change from Baseline in Simplified Disease Activity Index (SDAI) at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 24
- Change from Baseline in mTSS at Week 52
- The proportion of subjects with no radiographic progression from Baseline at Week 24 and 52
- Absolute value and change from Baseline in SF-36, FACIT-Fatigue, and the EQ-5D over time at Weeks 4, 12, 24 (except for SF-36 PCS and FACIT-Fatigue) and 52, and over time from Day 1 through Week 52
- Absolute value and change from Baseline in WPAI-RA at Weeks 4, 12, 24, and 52, and over time from Day 1 through Week 52

6.4.1. Definition of Other Secondary Efficacy Endpoints

6.4.1.1. ACR50 and ACR70

ACR50 and ACR70 are similarly defined as ACR20 (see Section 6.2.1), except that the improvement threshold from Baseline is 50% and 70%, respectively.

6.4.1.2. SDAI and CDAI

Simplified Disease Activity Index (SDAI) is a composite measure that sums the TJC28, SJC28, the SGA on a 0-10 scale, the PGA on a 0-10 scale, and the hsCRP (in mg/dL). SDAI is scored as follows {Aletaha 2005}:

$$\text{SDAI} = \text{TJC28} + \text{SJC28} + \text{SGA} + \text{PGA} + \text{CRP}$$

Higher SDAI score indicates more severe disease activity status.

Clinical Disease Activity Index (CDAI) is a further simplification of the SDAI that excludes the CRP, which is calculated using the following formula {Aletaha 2005}:

$$\text{CDAI} = \text{TJC28} + \text{SJC28} + \text{SGA} + \text{PGA}$$

CDAI can range from 0 to 76, with higher score indicating more severe disease activity status.

No component-level imputation will be performed for the calculation of both SDAI and CDAI. If any components are missing, the SDAI and CDAI will be set as missing.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis
Version 1.0

6.4.1.3. ACR-N

ACR-N is defined as the smallest percentage improvement from baseline in swollen joints, tender joints and the median of the following 5 items (PGA, SGA, subject's pain assessment, HAQ-DI score and CRP). It has a range between 0 and 100%. In particular,

\[ \text{ACR-N} = \min \{ \text{improvement in TJC66 }\% \}, \text{improvement in SJC66 }\% \}, \text{median [improvement in SGA }\% \}, \text{improvement in PGA }\% \}, \text{improvement in pain assessment }\% \}, \text{improvement in HAQ-DI }\% \}, \text{improvement in hsCRP }\% \} \}

If this calculation results in a negative value, then the ACR-N is set to 0. If any components are missing, the ACR-N will be set as missing.

6.4.1.4. EULAR response

Subject’s response will be categorized according to the following table based on the DAS28(CRP).

<table>
<thead>
<tr>
<th>Table 6-2.</th>
<th>EULAR Response Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAS28(CRP) at Visit</td>
<td>DAS28(CRP) Improvement from Baseline</td>
</tr>
<tr>
<td></td>
<td>&gt; 1.2</td>
</tr>
<tr>
<td>≤ 3.2</td>
<td>Good Response</td>
</tr>
<tr>
<td>&gt; 3.2 and ≤ 5.1</td>
<td>Moderate Response</td>
</tr>
<tr>
<td>&gt; 5.1</td>
<td>Moderate Response</td>
</tr>
</tbody>
</table>

6.4.1.5. European Quality of Life 5 Dimensions - 5 Levels (EQ-5D-5L)

The EQ-5D-5L is a standardized measure of health status of the subject at the visit (same day) that provides a simple, generic measure of health for clinical and economic appraisal. The EQ-5D-5L consists of 2 components: a descriptive system of the subject’s health and a rating of his or her current health state using a 0 to 100 VAS. The descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The subject is asked to indicate his/her health state by ticking (or placing a cross) in the box associated with the most appropriate statement in each of the 5 dimensions. The VAS records the subject’s self-rated health on a vertical VAS in which the endpoints are labeled ‘best imaginable health state’ on the top and ‘worst imaginable health state’ on the bottom. Higher EQ VAS indicates better health. The EQ-5D-5L will be scored according to the developer’s instructions (scoring guidelines).
6.4.1.6. Work Productivity and Activity Impairment questionnaire (WPAI) RA

The WPAI is a questionnaire developed to measure impairments in work activities in subjects with RA. The questionnaire consists of 6 questions:

Q1: currently employed,
Q2: work time missed due to RA,
Q3: work time missed due to other reasons,
Q4: hours actually worked,
Q5: degree RA affected productivity while working (0-10 VAS; with 0 indicating no effect and 10 indicating RA completely prevented the subject from working),
Q6: degree RA affected productivity in regular unpaid activities (0-10 VAS; with 0 indicating no effect and 10 indicating RA completely prevented the subject’s daily activities).

The recall period for questions 2 to 6 is 7 days. WPAI-RA outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, that is, worse outcomes, as follows:

- Absenteeism (work time missed) due to RA: 100B(Q2/(Q2+Q4))
- Presenteeism (impairment while working) due to RA: 100B(Q5/10)
- Work productivity loss (overall work impairment) due to RA: 100B(Q2/(Q2+Q4) + [(1-Q2/(Q2+Q4)) B (Q5/10)])
- Activity impairment due to RA: 100B(Q6/10)

If Question 1 (Are you currently employed?) is NO, then only the activity impairment score can be determined.

6.4.1.7. Healthcare Resource Utilization Questionnaire

The Healthcare Resource Utilization Questionnaire (HRUQ) is designed to assess healthcare usage during the previous three months across a number of direct medical cost domains.
6.4.1.8. Treatment Satisfaction Questionnaire for Medication (TSQM)

TSQM Scale scores are computed by adding the items loading on each factor. The lowest possible score is subtracted from this composite score and divided by the greatest possible score minus the lowest possible score. This provides a transformed score between 0 and 1 that should be multiplied by 100. Note that only one item may be missing from each scale before the subscale should be considered invalid for that respondent. Detailed scoring method is available in the Quintiles TSQM Scoring Manual v1.4.

6.4.1.9. Exploratory PRO

Subjects will be asked to rate the effect of their RA on their sexual functioning using a 0-100 VAS, with 0 indicating ‘rheumatoid arthritis has no effect on my sexual function’ and 100 indicating ‘rheumatoid arthritis completely inhibits my sexual function’.

6.4.2. Analysis Methods for Other Secondary Efficacy Endpoints

The FAS will be used for all summaries and analyses of other secondary efficacy endpoints. Statistical testing will be performed up to Week 24. The nominal p-values will be presented, if applicable. No formal statistical testing will be performed after Week 24, unless otherwise specified.

The proportion of subjects who achieve ACR20/50/70, and proportion of subjects with change in HAQ-DI ≥ 0.22 (ie, reduction in HAQ-DI ≥ 0.22), DAS28(CRP) ≥ 3.2 or DAS28(CRP) < 2.6, will be analyzed using the same logistic regression method with NRI as the primary endpoint analysis described in Section 6.2.3. Comparison will be made between each filgotinib dose group and placebo group. The 2-sided 95% CIs for the proportion based on normal approximation with the sample variance will be provided for each treatment group and visit. In addition, non-stratified response rate difference along with its 95% CI calculated based on the normal approximation with the pooled sample variance will be provided. The OC and LOCF values will be analyzed as a sensitivity approach to the primary method that uses NRI.

The proportion of subjects having no radiographic progression as defined by change from baseline in mTSS ≥ 0.5 at Week 24 will be analyzed using the same logistic regression method based on the OC. Comparison will be made between each filgotinib dose group and placebo group. The 2 sided 95% CIs for the proportion based on normal approximation with the sample variance will be provided for each treatment group. In addition, non-stratified response rate difference along with its 95% CI calculated based on the normal approximation with the pooled sample variance will be provided. The no radiographic progression as defined by change from baseline in mTSS ≥ 0.5, and the smallest detectable change (SDC) at Week 24 will also be analyzed using the logistic regression method based on the OC. SDC is the smallest detectable change computed from the variability in change from baseline in mTSS of the two imaging assessors {Bruynesteyn 2005}.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis Version 1.0

The change from Baseline in DAS28(CRP), CDAI, SDAI and mTSS will be analyzed using MMRM method that includes data at postbaseline visits up to the time point of interest. The MMRM models will include baseline value, stratification factors, treatment, visit, and treatment by visit interaction as fixed effects and subject being the random effect. The MMRM model will include all treatment groups. An unstructured variance-covariance matrix will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. Missing change scores due to missing study visits or early withdrawal will not be otherwise imputed using the MMRM approach. The LS means and 95% CI of the difference in mean change scores from Baseline in DAS28(CRP), CDAI, SDAI and mTSS between each filgotinib dose group and placebo group from the MMRM model will be provided for each postbaseline visit. Descriptive statistics of actual and change in DAS28(CRP), CDAI, SDAI and mTSS from baseline to postbaseline visits will also be provided by treatment group and visit.

Actual values and change from Baseline in individual components of the ACR response (including TJC68, SJC66, SGA, PGA, Subject’s pain assessment, and hsCRP), TJC28 and SJC28 will be summarized using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment group and visit. In addition, the MMRM model with baseline value, stratification factors, treatment, visit and treatment and visit interaction as fixed effects and subject being the random effect will be performed to compare each filgotinib dose group to placebo control group. The LS means and 95% CI of the difference between each filgotinib dose group and placebo control group will be provided.

Actual values of ACR-N will by summarize using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum, and maximum) by treatment and visit.

Number and percentage of subjects for EULAR responses will be presented for each visit by treatment.

For HRQoL endpoints, the analysis methods are detailed below:

- **SF-36**

  The transformed scores and change from baseline in 8 domains and 2 summary component scores (PCS and MCS) of the SF-36 will be summarized at each visit by treatment group using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum and maximum).
Filgotinib
gs-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis
Version 1.0

The change from Baseline in PCS and MCS will be analyzed using MMRM method that includes data at postbaseline visits up to the time point of interest, without imputation for missing data. The MMRM models will be used to evaluate treatment effect on change score from Baseline, with baseline value, stratification factors, treatment, visit, and treatment by visit interaction included as fixed effects and subject being the random effect. An unstructured variance-covariance matrix will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. To test for a treatment difference between each filgotinib dose group and placebo, the LS means along with the 95% CIs of the difference from MMRM, and p-value will be provided for each postbaseline visit.

- **FACIT -Fatigue**

  The change from baseline in FACIT-Fatigue scale score will be analyzed using similar MMRM models as for SF-36. To test for a treatment difference between each filgotinib dose group and placebo, LS mean difference along with 95% CI and p-values will be presented for each postbaseline visit. Missing change scores due to missing study visits or early withdrawal will not be otherwise imputed using the MMRM approach. Summary statistics of the actual score and change from baseline in FACIT-Fatigue scale score will also be displayed by treatment and visit.

- **EQ-5D-5L**

  Summary statistics of two outcomes of the EQ-5D-5L described as the following will be provided:

  — A health profile: the number and percentage of subjects at each categorical response for the 5 dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression) will be provided by treatment and visit.

  — A self-perceived current health score: calculated from subject level EQ VAS responses (continuous variable). Descriptive statistics (sample size, mean, SD, min, median, and max) will be provided by treatment and visit.

  The change from baseline in EQ VAS score will be analyzed using similar MMRM models as for SF-36. To test for a treatment difference between each filgotinib dose group and placebo, LS mean difference, 95% CI and p-values will be presented. Missing change scores due to missing study visits or early withdrawal will not be otherwise imputed using the MMRM approach.

- **WPAI - RA**

  Summary statistics of the actual scores and change from baseline in each of the four types of scores of WPAI-RA (absenteeism, presenteeism, work productivity loss, and activity impairment) will be summarized by treatment and visit using descriptive statistics (sample size, mean, SD, median, Q1, Q3, minimum and maximum).
The analyses of absenteeism, presenteeism, and work productivity loss will be based on subjects employed at both baseline and postbaseline visits.

- **HRUQ**
  Summary statistics will be provided for all the following parameters by treatment group and visit.
  
  *Healthcare resource utilization - Outpatient visits:*
  - Number and percentage of subjects with any medical visits, and the number of visits related to RA will be provided for the categories:
    - Outpatient Healthcare Provider
    - Emergency Room
    - Chiropractor, Physical or Occupational Therapist
  
  *Healthcare resource utilization - Inpatient hospitalization:*
  - Number and percentage of subject with hospital stays, and number of stays related to RA
  - Descriptive statistics for the total number of days in hospital related to RA
  - Number and percentage of subject with nursing home or rehabilitation center stays, and number of stays related to RA
  - Descriptive statistics for the total number of days in nursing home or rehabilitation center related to RA

- **TSQM**
  Descriptive statistics of the scores from each of the 4 domains (Effectiveness, Side Effects, Convenience, and Global Satisfaction) will be presented by treatment group and by visit.

- **Exploratory PRO**
  Summary statistics of the actual score and change from baseline will be displayed by treatment and visit. Subjects who marked “prefer not to answer” will not be included in this analysis.

Plots of proportions of subjects for categorical endpoints and mean ± SD for continuous endpoints (including each individual component of the ACR response) will be presented over time by treatment.

Additional statistical tests will be performed for the following efficacy endpoints;

- Compare each filgotinib dose group with adalimumab group at week 24 and 52 in proportion of subjects achieving ACR20, DAS28(CRP) ≤ 3.2 and DAS28(CRP) < 2.6 using similar logistic regression as described in Section 6.2.3, and in change from baseline in HAQ-DI, SF-36 PCS and FACIT-Fatigue using similar MMRM model as described in Section 6.3.2.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Version 1.0

- Compare filgotinib 200 mg group with placebo switching to filgotinib 200 mg group, and filgotinib 100 mg group with placebo switching to filgotinib 100 mg group at Week 52 in change from baseline in mTSS using similar MMRM model as described in Section 6.3.2.

- Compare filgotinib 200 mg group with placebo switching to filgotinib 200 mg group, and filgotinib 100 mg group with placebo switching to filgotinib 100 mg group at Week 52 in the proportion of subjects having no radiographic progression using similar logistic regression method as described in Section 6.4.2.

6.5. Changes From Protocol -Specified Efficacy Analyses

The following changes were made from the protocol-specified efficacy analyses:

- It has been observed from the blinded data that the enrollment rate is low in some combinations of stratification factors, which will result in the small cell counts for the stratified Cochran-Mantel-Haenszel test adjusted for the stratification factors. The logistic regression analysis with treatment and stratification factors in the model will be used instead to analyze binary endpoints.

- The following endpoints have been updated as key secondary efficacy endpoints in Section 6.3.1 and included in the hierarchy testing in Section 3.5, in order to support inclusion in the label.
  - Change from baseline in SF-36 physical component summary (PCS) at Week 12
  - Change from baseline in FACIT-Fatigue at Week 12

- Given the importance of determining the lowest effective dose for filgotinib in RA, the hierarchical ordering has been updated in Section 3.5 to test filgotinib 100 mg immediately after each test for filgotinib 200 mg (including the primary endpoint and each of the key secondary endpoints).

- Given the importance of achieving very low disease activity DAS28(CRP) < 2.6 in RA therapy, the testing for comparison of filgotinib with placebo on DAS28(CRP) H3.2 is removed from the hierarchy list, and the testing for comparison of filgotinib with adalimumab on DAS28(CRP) < 2.6 at week 12 is added to the hierarchy list. In addition, the comparison of filgotinib with placebo on DAS28(CRP) < 2.6 is evaluated at Week 12 prior to the specified non-responder assessment and start of standard of care therapy at Week 14.
7. SAFETY ANALYSES

7.1. Adverse Events and Deaths

7.1.1. Adverse Event Dictionary

Clinical and laboratory adverse events (AEs) will be coded using the MedDRA 21.0. System organ class (SOC), high-level term (HLT), preferred term (PT), and lower-level term (LLT) will be provided in the AE dataset.

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to toxicity criteria specified in the protocol. The severity grade of events for which the investigator did not record severity will be categorized as ‘missing’ for tabular summaries and data listings. The missing category will be listed last in summary presentation.

7.1.3. Relationship of Adverse Events to Study Drug

Related AEs are those for which the investigator selected ‘Related’ on the AE CRF to the question of ‘Related to Study Treatment’. Relatedness will always default to the investigator’s choice, not that of the Medical Monitor. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-subject data listings will show the relationship as missing.

7.1.4. Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if AEs met the definitions of SAE that were specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Pharmacovigilance & Epidemiology Department (PVE) before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as one or both of the following:

- Any AE with an onset date on or after the study drug start date and no later than 30 days after permanent discontinuation of study drug

- Any AE leading to premature discontinuation of study drug

For placebo subjects who were then re-randomized, TEAEs with an onset date prior to the first dose date of filgotinib will be allocated to the placebo controlled period. TEAEs with an onset date on or after the first dose date of filgotinib will be allocated to the re-randomized period.
7.1.5.2. Incomplete Dates

If the onset date of the AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset is the same as or after the month and year (or year) of the first dosing date of study drug, and
- The AE onset date is the same as or before the month and year (or year) of the date corresponding to 30 days after the date of the last dose of study drug

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent.

For placebo subjects who were then re-randomized, an AE with incomplete onset date that is prior to the month and year (or year) of the first dose date of filgotinib will be allocated to the placebo controlled period. An AE with incomplete onset date that is after the month and year (or year) of the first dose date of filgotinib will be allocated to the re-randomized period.

In addition, an AE with completely missing onset date or incomplete onset date that is same as the month and year (or year) of the first dose date of filgotinib, and stop date that is prior to the date (or month and year if day is not recorded; or year alone if month is not recorded) of the first dose of filgotinib will be allocated to the placebo controlled period. An AE with completely missing onset date or incomplete onset date that is same as the month and year (or year) of the first dose date of filgotinib, and stop date that is missing, or the same as, or after the date (or month and year if day is not recorded; or year alone if month is not recorded) of the first dose of filgotinib will be allocated to the re-randomized period.

7.1.6. Summaries of Adverse Events and Deaths

Treatment-emergent AEs will be summarized based on the Safety Analysis Set.

7.1.6.1. Summaries of AE incidence in Combined Severity Grade Subsets

The number and percentage of subjects who experienced at least 1 TEAE will be provided and summarized by SOC, HLT, PT, and treatment group. For other AEs described below, summaries will be provided by SOC, PT, and treatment group:

- TEAEs of Grade 3 or higher (by maximum severity)
Filgotinib

GS-US-417-0301

Statistical Analysis Plan - Week 24 Analysis and Final Analysis

Version 1.0

- TEAEs of Grade 2 or higher
- All TE treatment-related AEs
- TE Treatment-related AEs of Grade 3 or higher (by maximum severity)
- TE Treatment-related AEs of Grade 2 or higher
- All TE SAEs
- All TE treatment-related SAEs
- All TEAEs leading to premature discontinuation of any study drug
- All TEAEs leading to premature discontinuation of study
- All TE SAEs leading to death (ie, outcome of death)
- All TEAEs leading to temporary interruption of any study drug

A brief, high-level summary of AEs described above will be provided by treatment group and by the number and percentage of subjects who experienced the above AEs. All deaths observed in the study will be also included in this summary.

Multiple events will be counted only once per subject in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and HLT within each SOC (if applicable), and then by PT in descending order of total frequency within each SOC. For summaries by severity grade, the most severe grade will be used for those AEs that occurred more than once in an individual subject during the study.

In addition to the above summary tables, all TEAEs and TE treatment-related AEs will be summarized by PT only, in descending order of total frequency.

In addition, data listings will be provided for the following:

- All AEs, indicating whether the event is treatment emergent
- All AEs of Grade 3 or higher
- All AEs of Grade 2 or higher
- SAEs
- Deaths
- All SAEs leading to death (ie, outcome of death)
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

- AEs leading to premature discontinuation of any study drug
- AEs leading to premature discontinuation of study
- AEs leading to temporary interruption of any study drug

7.1.7. Adverse Events of Special Interest

Events of interest will be identified by the use of either SMQs or MSTs. However, should additional cases not detected by the predefined search term listings be identified during the clinical review process, these cases will also be reported by respective category.

7.1.7.1. Adjudication Committee for MACE

An independent cardiovascular safety endpoint adjudication committee (CVEAC) will be formed to periodically review and adjudicate all potential major adverse cardiovascular events (MACE) events. MACE events are defined as cardiovascular death, non-fatal myocardial infarction and non-fatal stroke.

To identify the MACE, the following potential cases identified using MedDRA search term (MST) searches will be adjudicated:

- All deaths
- Cardiovascular events (meeting serious criteria)
- Myocardial infarction (Narrow)
- Hospitalization for unstable angina
- Transient ischemic attack
- Stroke
- Hospitalization for cardiac failure
- Percutaneous coronary intervention

The CVEAC will review those potential MACE and related clinical data to determine whether a MACE has developed. The CVEAC’s role and responsibilities and the data to be provided to the CVEAC are described in a mutually agreed upon CVEAC charter. The CVEAC charter defines the CVEAC membership, adjudication process, meeting logistics, and meeting frequency.

The number and percentage of subjects with positively adjudicated MACE will be summarized by treatment group using the PT.

A by-subject listing for all subjects who have potential MACE and who have a positively adjudicated MACE at any time will be provided.
7.1.7.2. Other Adverse Events of Special Interest

In addition to general safety parameters and MACE, safety information on other adverse events of special interest (AESIs) will also be analyzed. AESIs will be identified by laboratory results, standardized MedDRA queries (SMQs), sponsor defined MSTs, or a combination of these methods as indicated below.

- All infections (defined as all PTs in the Infections and Infestations SOC)
- Serious infections (defined as all PTs in the Infections and Infestations SOC that are SAEs)
- Infections of special interest as defined below
  a) Herpes zoster
  b) Active tuberculosis
  c) Opportunistic infections
  d) Hepatitis B or C infections
- Deep vein thrombosis (DVT) and pulmonary embolism (PE)
- Malignancy (including lymphoma; not including nonmelanoma skin cancer)
- Nonmelanoma skin cancer
- Gastrointestinal (GI) perforations

The number and percentage of subjects with aforementioned events of special interest will be provided by the PT for each AE of special interests.

A by-subject listing for all subjects having AE of special interests at any time will be provided for each AE of special interest.

7.1.8. Additional Analysis for Adverse Events

The number and proportion of subjects who experienced at least one treatment-emergent SAE, or one treatment-emergent AE of special interests will be summarized by treatment group. The comparisons of the proportions between the placebo group during placebo controlled period and the filgotinib treatment groups will be performed. The 2-sided 95% CI of the proportion based on normal approximation with the sample variance will be provided for each treatment group. In addition, the difference in proportions between treatment groups with its 95% CI based on the normal approximation with the pooled sample variance will be provided.
7.2. Laboratory Evaluations

Laboratory data collected during the study will be analyzed and summarized using both quantitative and qualitative methods. Summaries of laboratory data will be provided for the Safety Analysis Set and will include data collected up to the last dose of any study drug plus 30 days for subjects who have permanently discontinued study drug, or all available data at the time of the database snapshot for subjects who were ongoing at the time of an interim analysis. The analysis will be based on values reported in conventional units. When values are below the LOQ, they will be listed as such, and the closest imputed value will be used for the purpose of calculating summary statistics as specified in Section 3.7.

A baseline laboratory value will be defined as the last nonmissing measurement obtained on or prior to the date/time of first dose of any study drug.

For placebo subjects who were then re-randomized, the baseline laboratory value for the re-randomization period will be the last available measurement obtained on or prior to the date/time of first dose of filgotinib. The lab abnormalities after the first dose of filgotinib will be assessed from the baseline of the re-randomization period.

A by-subject listing for laboratory test results will be provided by subject ID number and visit in chronological order for hematology, serum chemistry, and urinalysis separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the CTCAE severity grade will be flagged in the data listings, as appropriate.

No formal statistical testing is planned.

7.2.1. Summaries of Numeric Laboratory Results

Descriptive statistics of Baseline values, values at each postbaseline visit and change from Baseline at each postbaseline visit will be provided by treatment group for the following laboratory tests:

- Hematology
  - Hematocrit
  - Hemoglobin
  - Platelet count
  - Red blood cell count
  - White blood cell (WBC) count
  - Mean corpuscular volume
  - Lymphocytes
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

Version 1.0

- Monocytes
- Neutrophils
- Eosinophils
- Basophils

- Chemistry
  - Alanine aminotransferase (ALT)
  - Aspartate aminotransferase (AST)
  - Alkaline phosphatase (ALP)
  - Total bilirubin
  - Serum creatinine
  - Creatinine clearance by Cockcroft-Gault formula
  - Creatinine phosphokinase (CPK)
  - Glucose

- Lipid
  - Triglycerides
  - Total cholesterol
  - HDL
  - LDL
  - LDL/HDL ratio

Change from baseline to a postbaseline visit will be defined as the visit value minus the baseline value. The mean, median, Q1, Q3, minimum, and maximum values will be displayed to the reported number of digits; SD values will be displayed to the reported number of digits plus 1.

Median (Q1, Q3) of the observed values for the laboratory tests specified above will be plotted using a line plot by treatment group and visit.
In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3.

7.2.2. Graded Laboratory Value

The CTCAE Version 4.03 will be used to assign toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (i.e., increased, decreased) will be presented separately.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent laboratory abnormalities are defined as values that increase at least 1 toxicity grade from baseline at any postbaseline time point, up to and including the date of last dose of any study drug plus 30 days for subjects who permanently discontinued study drug, or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.

For placebo subjects who were then re-randomized, treatment-emergent laboratory abnormalities of placebo period are defined as values that increase at least 1 toxicity grade from the baseline at any postbaseline time point, up to and including the first dose date of filgotinib. Treatment-emergent laboratory abnormalities of re-randomized period are defined as values that increase at least 1 toxicity grade from the new baseline after Week 24 at any postbaseline time point, up to and including the date of last dose of any study drug plus 30 days, or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.

Placebo subjects who were not re-randomized have treatment-emergent laboratory abnormalities in the placebo period which are defined as values that increase 1 toxicity grade from the baseline at any postbaseline time point, up to and including the date of last dose of any study drug plus 30 days.

If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent.

7.2.2.2. Treatment-Emergent Marked Laboratory Abnormalities

Treatment-emergent marked laboratory abnormalities are defined as values that increase from baseline by at least 3 toxicity grades at any postbaseline time point, up to and including the date of the last dose of any study drug plus 30 days for subjects who permanently discontinued study drug or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis

For placebo subjects who were then re-randomized, treatment-emergent marked laboratory abnormalities of placebo period are defined as values that increase from baseline by at least 3 toxicity grades at any postbaseline time point, up to and including the first dose date of filgotinib. Treatment-emergent marked laboratory abnormalities of re-randomized period are defined as values that increase from the new baseline after Week 24 by at least 3 toxicity grades at any postbaseline time point, up to and including the date of last dose of any study drug plus 30 days, or the last available date in the database snapshot for subjects who were still on treatment at the time of an interim analysis.

Placebo subjects who were not re-randomized have treatment-emergent marked laboratory abnormalities in the placebo period which are defined as values that increase from baseline by at least 3 toxicity grades at any postbaseline time point, up to and including the date of last dose of any study drug plus 30 days.

If the relevant baseline laboratory value is missing, any Grade 3 or higher values observed within the time frame specified above will be considered treatment-emergent marked abnormalities.

7.2.2.3. Summaries of Laboratory Abnormalities

Laboratory data that are categorical will be summarized using the number and percentage of subjects in the study with the given response at baseline and each scheduled postbaseline visit.

The following summaries (number and percentage of subjects) for treatment-emergent laboratory abnormalities will be provided by laboratory test and treatment group: subjects will be categorized according to the most severe postbaseline abnormality grade for a given laboratory test:

- Graded laboratory abnormalities
- Grade 3 or higher laboratory abnormalities
- Marked laboratory abnormalities

For all summaries of laboratory abnormalities, the denominator is the number of subjects with nonmissing postbaseline values up to 30 days after last dosing date.

A by-subject listing of treatment-emergent Grade 3 or higher laboratory abnormalities and marked laboratory abnormalities will be provided separately by subject ID number and visit in chronological order. These listings will include all test results that were collected throughout the study for the laboratory test of interest, with all applicable severity grades or abnormal flags displayed.
7.2.3. Laboratory Evaluations of Special Interest

7.2.3.1. Liver-Related Laboratory Evaluations

Liver-related abnormalities after initial study drug dosing will be examined and summarized using the number and percentage of subjects who were reported to have the following laboratory test values for postbaseline measurements:

- AST: (a) > 3 times the upper limit of reference range (ULN); (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- ALT: (a) > 3 x ULN; (b) > 5 x ULN; (c) > 10 x ULN; (d) > 20 x ULN
- AST or ALT > 3 x ULN and total bilirubin > 2 x ULN

The summary will include data from all postbaseline visits up to 30 days after the last dose of any study drug. For individual laboratory tests, subjects will be counted once based on the most severe postbaseline values. For both the composite endpoints of AST or ALT and total bilirubin, subjects will be counted once when the criteria are met at the same postbaseline visit date. The denominator is the number of subjects in the Safety Analysis Set who have nonmissing postbaseline values of all relevant tests at the same postbaseline visit date. A listing of subjects who met at least 1 of the above criteria will be provided.

7.2.3.2. Complete Blood Count-Related Laboratory Evaluations

Complete blood count (CBC)-related abnormalities such as anemia, leucopenia, neutropenia, lymphopenia, and thrombocytopenia after initial study drug dosing will be examined and summarized using the number and percentage of subjects who were reported to have the following laboratory test values for postbaseline measurements:

- Hemoglobin: (a) any postbaseline worsening CTCAE grade from baseline; (b) baseline value of less than Grade 3 and increase to Grade 3 or higher at worst postbaseline; (c) baseline value of less than Grade 3 and increase to Grade 4 at worst postbaseline
- WBC count: (a) any postbaseline worsening CTCAE grade from baseline; (b) baseline value of less than Grade 3 and increase to Grade 3 or higher at worst postbaseline; (c) baseline value of less than Grade 3 and increase to Grade 4 at worst postbaseline
- Absolute neutrophil count: (a) any postbaseline worsening CTCAE grade from baseline; (b) baseline value of less than Grade 3 and increase to Grade 3 or higher at worst postbaseline; (c) baseline value of less than Grade 3 and increase to Grade 4 at worst postbaseline
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis

- Lymphocyte count: (a) any postbaseline worsening CTCAE grade from baseline; (b) baseline value of less than Grade 3 and increase to Grade 3 or higher at worst postbaseline; (c) baseline value of less than Grade 3 and increase to Grade 4 at worst postbaseline

- Platelet count: (a) any postbaseline worsening CTCAE grade from baseline; (b) baseline value of less than Grade 3 and increase to Grade 3 or higher at worst postbaseline; (c) baseline value of less than Grade 3 and increase to Grade 4 at worst postbaseline

The summary will include data from all postbaseline visits up to 30 days after the last dose of any study drug.

7.3. **Body Weight and Vital Signs**

Descriptive statistics will be provided by treatment group for body weight, BMI and vital signs (systolic and diastolic blood pressures [mmHg], pulse [beats/min]) as follows:

- Baseline value
- Values at each postbaseline visit
- Change from baseline at each postbaseline visit

A baseline value will be defined as the last available value collected on or prior to the date/time of first dose of any study drug.

For placebo subjects who were then re-randomized, the baseline value for re-randomization period will be defined as the last available value collected on or prior to the date/time of the first dose of filgotinib.

Change from baseline to a postbaseline visit will be defined as the postbaseline value minus the baseline value. Body weight and vital signs measured at unscheduled visits will be included for the baseline value selection.

In the case of multiple values in an analysis window, data will be selected for analysis as described in Section 3.8.3. No formal statistical testing is planned.

A by-subject listing of vital signs (systolic and diastolic blood pressure [mmHg], pulse [beats/min], respiration [breaths/min], and body temperature [°C]) will be provided by subject ID number and visit in chronological order. In the same manner, a by-subject listing of body weight, height, and BMI will be provided separately.

7.4. **Prior and Concomitant Medications**

Medications collected at screening and during the study will be coded using the World Health Organization (WHO) Drug dictionary version BSEP17.

All the analyses in this section will be performed for general prior /concomitant medications and RA-specific prior/concomitant medications separately, unless otherwise specified.
7.4.1. Prior Medications

Prior medications are defined as any medication taken before a subject took the first study drug.

Prior medications will be summarized by Anatomical Therapeutic Chemical (ATC) drug class preferred name using the number and percentage of subjects for each treatment group and overall. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be provided by preferred name in order of descending overall frequency. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to the first dosing date of any study drug will be included in the prior medication summary regardless of when the stop date is. If a partial start date is entered the medication will be considered prior unless the month and year (if day is missing) or year (if day and month are missing) of the start date are after the first dosing date. Medications with a completely missing start date will be included in the prior medication summary, unless otherwise specified.

Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

7.4.2. Concomitant Medications

Concomitant medications are defined as medications taken while a subject took study drug. Use of concomitant medications will be summarized by ATC drug class preferred name using the number and percentage of subjects for each treatment group. A subject reporting the same medication more than once will be counted only once when calculating the number and percentage of subjects who received that medication. The summary will be provided by preferred term in descending overall frequency. For drugs with the same frequency, sorting will be done alphabetically.

For the purposes of analysis, any medication with a start date prior to or on the first dosing date of any study drug and continued to take after the first dosing date, or started after the first dosing date but prior to or on the last dosing date of study drug will be considered concomitant medications. Medications started and stopped on the same day as the first dosing date or the last dosing date of any study drug will also be considered concomitant. Medications stopped on the same day as the first dosing date will be considered concomitant. Medications with a stop date prior to the date of first dosing date of any study drug or a start date after the last dosing date of any study drug will be excluded from the concomitant medication summary. If a partial stop date is entered, any medication with the month and year (if day is missing) or year (if day and month are missing) prior to the date of first study drug administration will be excluded from the concomitant medication summary. If a partial start date is entered, any medication with the month and year (if day is missing) or year (if day and month are missing) after the study drug stop date will be excluded from the concomitant medication summary. Medications with completely missing start and stop dates will be included in the concomitant medication summary, unless otherwise specified.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan – Week 24 Analysis and Final Analysis

For placebo subjects who were then re-randomized, concomitant medication with a start date prior to the first dose date of filgotinib will be allocated to the placebo controlled period. Concomitant medication with a start date on or after first dose date of filgotinib will be allocated to the re-randomized period.

For placebo subjects who were then re-randomized, a concomitant medication with incomplete start date that is prior to the month and year (or year) of the first dose date of filgotinib will be allocated to the placebo controlled period. A concomitant medication with incomplete start date that is after the month and year (or year) of the first dose date of filgotinib will be allocated to the re-randomized period.

In addition, an concomitant medication with completely missing start date or incomplete start date that is same as the month and year (or year) of the first dose date of filgotinib, and stop date that is prior to the date (or month and year if day is not recorded; or year alone if month is not recorded) of the first dose of filgotinib will be allocated to the placebo controlled period. A concomitant medication with completely missing start date or incomplete start date that is same as the month and year (or year) of the first dose date of filgotinib, and stop date that is missing, or the same as, or after the date (or month and year if day is not recorded; or year alone if month is not recorded) of the first dose of filgotinib will be allocated to the re-randomized period.

Summaries will be based on the Safety Analysis Set. No formal statistical testing is planned.

All prior and concomitant medications (other than per-protocol study drugs) will be provided in a by-subject listing sorted by subject ID number and administration date in chronological order.

7.5. Electrocardiogram Results

7.5.1. Investigator Electrocardiogram Assessment

A shift table of the investigators’ assessment of ECG results at each visit compared with baseline values will be presented by treatment group using the following categories: normal; abnormal (not clinically significant); abnormal (clinically significant); or missing. The number and percentage of subjects in each cross-classification group of the shift table will be presented. Subjects with a missing value at baseline or postbaseline will not be included in the denominator for percentage calculation.

No formal statistical testing is planned.

A by-subject listing for ECG assessment results will be provided by subject ID number and visit in chronological order.

7.6. Other Safety Measures

A data listing will be provided for subjects who become pregnant during the study.

7.7. Changes From Protocol -Specified Safety Analyses

There are no deviations from the protocol-specified safety analyses.
8. PHARMACOKINETIC ANALYSES

Concentrations of filgotinib and its metabolite GS-829845 in plasma will be determined using validated bioanalytical assays. Plasma PK parameters of filgotinib and GS-829845, will be evaluated.

8.1. PK Analyses Related to Intensive PK Sampling

Steady-state PK over a 24 hour dosing interval will be determined in subjects in the PK Substudy Analysis Set.

8.1.1. Estimation of Pharmacokinetic Parameters

PK parameters will be estimated using Phoenix WinNonlin® software using standard noncompartmental methods. The linear/log trapezoidal rule will be used in conjunction with the appropriate noncompartmental model, with input values for dose level, dosing time, plasma concentration, and corresponding real-time values, based on drug dosing times whenever possible.

All predose sample times before time-zero will be converted to 0. Predose samples may also serve as the 24-hr post dose sample if appropriate.

For area under the curve (AUC), samples BLQ of the bioanalytical assays occurring prior to the achievement of the first quantifiable concentration will be assigned a concentration value of 0 to prevent overestimation of the initial AUC. Samples that are BLQ at all other time points will be treated as missing data in WinNonlin. The nominal time point for a key event or dosing interval (-) may be used to permit direct calculation of AUC over specific time intervals. The appropriateness of this approach will be assessed by the PK scientist on a profile-by-profile basis.

Pharmacokinetic parameters such as AUC$_{\text{tau,inf}}$ and $t_{1/2}$ are dependent on an accurate estimation of the terminal elimination phase of drug. The appropriateness of calculating these parameters will be evaluated upon inspection of PK data on a profile-by-profile basis by the PK scientist.

8.1.2. Pharmacokinetic Parameters

PK parameters will be generated for all subjects in the PK Substudy analysis set. The analytes presented in Table 8-1 will be evaluated if data are available.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Study Treatments and Associated Analytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filgotinib 200 mg</td>
<td>Filgotinib, GS-829845</td>
</tr>
<tr>
<td>Filgotinib 100 mg</td>
<td>Filgotinib, GS-829845</td>
</tr>
</tbody>
</table>
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis V1.0

The analytes and parameters presented in Table 8-2 will be used to evaluate the PK objectives of the study. The PK parameters to be estimated in this study are listed and defined in the Pharmacokinetic Abbreviations section.

<table>
<thead>
<tr>
<th>Table 8-2. Pharmacokinetic Parameters for Each Analyte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyte</td>
</tr>
<tr>
<td>Filgotinib</td>
</tr>
<tr>
<td>GS-829845</td>
</tr>
</tbody>
</table>

In addition, molar ratio of metabolite to parent exposure (AUC_{tau} and C_{max}) may be calculated for individual subjects and summarized by treatment. The GS-829845 to filgotinib ratio for AUC_{tau} will be calculated by dividing the AUC_{tau} (in h*nmol/L) of GS-829845 by the AUC_{tau} (in h*nmol/L) of filgotinib. The GS-829845 to filgotinib ratio for C_{max} will be calculated by dividing the C_{max} (in nmol/L) of GS-829845 by the C_{max} (in nmol/L) of filgotinib.

8.1.3. Statistical Analysis Methods

Individual subject concentration data and individual subject PK parameters for filgotinib and GS-829845 will be listed and summarized using descriptive statistics by treatment. Summary statistics (number of subjects [n], mean, SD, coefficient of variation [%CV], median, minimum, maximum, Q1, and Q3) will be presented for both individual subject concentration data by time point and treatment; and individual subject PK parameters by treatment. Moreover, the geometric mean, 95% CI, and the mean and SD of the natural log-transformed values will be presented for individual subject PK parameter data.

Individual concentration data listings and summaries will include all subjects with concentration data. The sample size for each time point will be based on the number of subjects with nonmissing concentration data at that time point. The number of subjects with concentration BLQ will be presented for each time point. For summary statistics, BLQ values will be treated as 0 at predose and one-half of the lower limits of quantitation (LLOQ) for postdose time points.

Individual PK parameter data listings and summaries will include all subjects for whom PK parameter(s) can be derived. The sample size for each PK parameter will be based on the number of subjects with nonmissing data for that PK parameter.

The following tables may be provided for each analyte by treatment:

- Individual subject concentration data and summary statistics
- Individual subject plasma PK parameters and summary statistics.
Filgotinib  
GS-US-417-0301  
Statistical Analysis Plan - Week 24 Analysis and Final Analysis  
Version 1.0

The following figures may be provided for each analyte by treatment (for PK substudy only):

- Individual subject concentration data versus time (on linear and semilogarithmic scales)
- Mean (± SD) concentration data versus time (on linear and semilogarithmic scales)
- Median (Q1, Q3) concentration data versus time (on linear and semilogarithmic scales)

Individual, mean, and median postdose concentration values that are H LLOQ will not be displayed in the figures and remaining points connected.

The following listings may be provided:

- PK sampling details (and PK concentrations) by subject, including procedures, differences in scheduled and actual draw times, and sample age
- Individual data on determination of plasma half-life and corresponding regression correlation coefficient.
9. **BIOMARKER ANALYSIS**

A separate biomarker analysis plan (BAP) will document methods to analyze biomarker assessments.
10. REFERENCES


Yuan Y. Sensitivity Analysis in Multiple Imputation for Missing Data. 2014:
11. SOFTWARE


nQuery Advisor(R) Version 4.0. Statistical Solutions, Cork, Ireland.
12. SAP REVISION

<table>
<thead>
<tr>
<th>Revision Date</th>
<th>Section</th>
<th>Summary of Revision</th>
<th>Reason for Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>27 Nov 2018</td>
<td>Throughout the document</td>
<td>Editorial changes</td>
<td>Clarification</td>
</tr>
<tr>
<td>27 Nov 2018</td>
<td>Section 3.5</td>
<td>Update hierarchy list</td>
<td>Reflect endpoints importance</td>
</tr>
<tr>
<td>27 Nov 2018</td>
<td>Section 6.1</td>
<td>Estimands</td>
<td>Add definition of estimands</td>
</tr>
<tr>
<td>27 Nov 2018</td>
<td>Section 6.3.3</td>
<td>Non-inferiority sensitivity analysis</td>
<td>Add MI and TP analysis for non-inferiority tests</td>
</tr>
<tr>
<td>15 Jan 2018</td>
<td>Throughout the document</td>
<td>Editorial changes</td>
<td>Clarification</td>
</tr>
</tbody>
</table>
13. APPENDIX

Appendix 1. Lists of RA Medications

1) List of bDMARDs and Investigational bDMARDs (WHO Preferred Terms)
   - ABATACEPT
   - ADALIMUMAB
   - CABIRALIZUMAB
   - CERTOLIZUMAB
   - CERTOLIZUMAB PEGOL
   - CLAZAKIZUMAB
   - DENOSUMAB
   - ETANERCEPT
   - GOLIMUMAB
   - INFLIXIMAB
   - INTERLEUKIN-2
   - SIRUKUMAB
   - TOCILIZUMAB

2) List of Oral Corticosteroids (WHO Preferred Terms)
   - BETAMETHASONE
   - BETAMETHASONE SODIUM PHOSPHATE
   - CORTISONE
   - DEXAMETHASONE
   - DEXAMETHASONE PALMITATE
   - DEXAMETHASONE PHOSPHATE
   - DEXAMETHASONE SODIUM PHOSPHATE
Filgotinib
GS-US-417-0301
Statistical Analysis Plan ~ Week 24 Analysis and Final Analysis Version 1.0

- MEPREDNISONE
- METHYL PREDNISOLONE
- METHYL PREDNISOLONE ACETATE
- METHYL PREDNISOLONE SODIUM Succinate
- MOMETASONE FURATE
- PREDNISOLONE
- PREDNISOLONE FARNESYLATE
- PREDNISOLONE
- STEROIDS

3) List of Anti-malarials (WHO Preferred Terms)
   - CHLOROQUINE
   - CHLOROQUINE PHOSPHATE
   - CHLOROQUINE SULFATE
   - HYDROXY CHLOROQUINE
   - HYDROXY CHLOROQUINE SULFATE
Appendix 2. Health Assessment Questionnaire Disability Index (HAQ-DI)

The HAQ-DI score is defined as the average of the scores of eight functional categories (dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities), usually administered by the subject. Responses in each functional category are collected as 0 (without any difficulty) to 3 (unable to do a task in that area), with or without aids or devices.

The highest score for questions in each category (range 0 to 3) determines the score for the category, unless aids or devices are required. Dependence on equipment or physical assistance increases a lower score (i.e., scores of 0 or 1) to the level of 2 to more accurately represent underlying disability. The eight category scores are averaged into an overall HAQ-DI score on a scale from 0 (no disability) to 3 (completely disabled) when 6 or more categories are non-missing. If more than 2 categories are missing, the HAQ-DI score is set to missing. The HAQ-DI can be treated as a continuous measure.

The HAQ-DI score using aids (and/or devices) is computed by taking the maximum score of the questions in each category (range: [0, 3]) and whether or not aids/devices are used (0 or 1):

\[
A = \max(\text{dressing & grooming category questions, } 2\times\text{aids indicator}) + \\
\max(\text{eating category questions, } 2\times\text{aids indicator}) + \\
\max(\text{walking category questions, } 2\times\text{aids indicator}) + \\
\max(\text{hygiene category questions, } 2\times\text{aids indicator}) + \\
\max(\text{reach category questions, } 2\times\text{aids indicator}) + \\
\max(\text{grip category questions, } 2\times\text{aids indicator}) + \\
\max(\text{usual activities category questions, } 2\times\text{aids indicator})
\]

\[
\text{HAQ-DI} = A / (\text{total number of categories with at least 6 non-missing})
\]

The following table shows the contribution of the 43 questions used to calculate the HAQ-DI:
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis
Version 1.0

<table>
<thead>
<tr>
<th>HAQ-DI Category:</th>
<th>Category Questions</th>
<th>Aids/Devices Indicators</th>
<th>HAQ-DI Category Score with Aids/Devices Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dressing / Grooming</td>
<td>HAQ0101, HAQ0102 (DRESS, HAIR)</td>
<td>HAQ0114, HAQ0119 (DRSG, GROOM)</td>
<td>Using each question with a scale of 0-3, calculate the category score as the maximum of the category questions.</td>
</tr>
<tr>
<td></td>
<td>Arising</td>
<td>HAQ0103, HAQ0104 (STAND, BED)</td>
<td>If the Aids/Devices indicator is 'No', no need to adjust the category score.</td>
</tr>
<tr>
<td></td>
<td>Eating</td>
<td>HAQ0105, HAQ0106, HAQ0107 (MEAT, LIFT, MILK)</td>
<td>If the Aids/Devices indicator is 'Yes' and the category score is &lt;2, then the category score with the Aids/Devices is set to 2.</td>
</tr>
<tr>
<td></td>
<td>Walking</td>
<td>HAQ0108, HAQ0109 (WALK, STEPS)</td>
<td>If the Aids/Devices indicator is 'Yes' and the category score is &gt;2, then the category score with the Aids/Devices is the calculated category score without adjustment.</td>
</tr>
<tr>
<td></td>
<td>Hygiene</td>
<td>HAQ0123, HAQ0124, HAQ0125 (WASH, BATH, TOILET)</td>
<td>For example: The Dressing/Grooming category score is 2 if subject answered 1 for both questions 1 and 2 and 'Yes' for both question 14 and 19.</td>
</tr>
<tr>
<td></td>
<td>Reach</td>
<td>HAQ0126, HAQ0127 (REACH, BEND)</td>
<td>In the HAQ-DI score calculation, questions on other device/aid will not be used.</td>
</tr>
<tr>
<td></td>
<td>Grip</td>
<td>HAQ0128, HAQ0129, HAQ0130 (OPENCAR, JAR, FAUCET)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Activity</td>
<td>HAQ0131, HAQ0132, HAQ0133 (SHOP, INCAR, CHORES)</td>
<td></td>
</tr>
</tbody>
</table>

**Handling Missing Data**: If no more than 2 categories have missing category scores, then the HAQ-DI is the mean of the non-missing category scores. Otherwise, the HAQ-DI score is set to missing.

If any of the category questions are missing, but the aids/device indicator is non-missing, the category score can still be computed. However, if all category questions and its aids/device indicators are missing, then the category score is considered missing.
Appendix 3. van der Heijde modified Total Sharp Score (mTSS)

1) Definition of mTSS

The joint erosion score is a summary of erosion severity in 32 joints of the hands and 12 joints in the feet, as shown below for the joints assessed for erosions.

<table>
<thead>
<tr>
<th>Hands/Fingers MCP:</th>
<th>Hands/Wrist:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metacarpophalangeal joints I − V (n = 10)</td>
<td>Multangular bones (trapezium &amp; trapezoid) (n = 2)</td>
</tr>
<tr>
<td>Hands/Fingers PIP:</td>
<td>Hands/Wrist:</td>
</tr>
<tr>
<td>Proximal interphalangeal joints II − V (n = 8)</td>
<td>Scaphoid (Navicular) bones (n = 2)</td>
</tr>
<tr>
<td>Hands/Fingers IP (thumbs):</td>
<td>Hands/Wrist:</td>
</tr>
<tr>
<td>Interphalangeal joints (n = 2)</td>
<td>Lunate bones (n = 2)</td>
</tr>
<tr>
<td>Hands/Fingers (thumbs):</td>
<td>Feet/Toes MTP:</td>
</tr>
<tr>
<td>Proximal Metacarpal joints (n = 2)</td>
<td>Metatarsophalangeal joints I − V (n = 10)</td>
</tr>
<tr>
<td>Hands/Wrist:</td>
<td>Feet/Toes IP (big toes):</td>
</tr>
<tr>
<td>Distal Radius (n = 2)</td>
<td>Interphalangeal joints (n = 2)</td>
</tr>
<tr>
<td>Hands/Wrist:</td>
<td></td>
</tr>
<tr>
<td>Distal Ulna (n = 2)</td>
<td></td>
</tr>
</tbody>
</table>

An erosion score of 0 to 5 is given to each joint in the hands and wrists, and a score of 0 to 10 is given to each joint in the feet. Each hand and wrist joint is scored, according to the surface area involved, from 0 to 5, with 5 indicating extensive loss of bone from more than one half of the articulating bone (0 indicates no erosion). Because each side of a foot joint is graded on this scale, the maximum erosion score for a foot joint is 10. Thus, the maximal erosion score is (32*5) + (12*10) = 280. A joint may not be evaluable due to surgery (ie, joint replacement) or may be radiographically insufficient for reading.

The joint space narrowing (JSN) score summarizes the severity of JSN in 30 joints of the hands/wrists and 12 joints of the feet, as shown below for the joints assessed for JSN.

<table>
<thead>
<tr>
<th>Hand/Fingers MCP:</th>
<th>Hand/Wrist:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metacarpophalangeal joints I − V (n = 10)</td>
<td>Capitate-scaphoid joint (n = 2)</td>
</tr>
<tr>
<td>Hand/Fingers PIP:</td>
<td>Hand/Wrist:</td>
</tr>
<tr>
<td>Proximal interphalangeal joints II − V (n = 8)</td>
<td>Scaphoid-radius joint (n = 2)</td>
</tr>
<tr>
<td>Hand/Fingers CMC:</td>
<td>Foot/Toes MTP:</td>
</tr>
<tr>
<td>Carpometacarpal joints III, IV, V (n = 6)</td>
<td>Metatarsophalangeal joints (n = 10)</td>
</tr>
<tr>
<td>Hand/Wrist:</td>
<td>Foot/Toes IP:</td>
</tr>
<tr>
<td>Scaphoid-Trapezium joint (n = 2)</td>
<td>Interphalangeal joints (n = 2)</td>
</tr>
</tbody>
</table>

Assessment of JSN for each hand (15 joints per hand) and foot (6 joints per foot), including subluxation, is scored from 0 to 4, with 0 indicating normal or no narrowing and 4 indicating complete loss of joint space, bony ankylosis, or luxation. Thus, the maximal JSN score is (30*4) + (12*4) = 168. A joint may not be evaluable due to subluxation, luxations, surgery (ie, joint replacement) or may be radiographically insufficient for reading.
The mTSS is defined as the erosion score plus the joint space narrowing score.

The maximum scores (adding up both hands/wrists and feet) are:

<table>
<thead>
<tr>
<th>Radiography</th>
<th>Maximum score of hands/wrists</th>
<th>Maximum score of feet</th>
<th>Maximum score (add up hands/wrists and feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosion score</td>
<td>32 joints * 5 = 160</td>
<td>12 joints * 10 = 120</td>
<td>280</td>
</tr>
<tr>
<td>JSN score</td>
<td>30 joints * 4 = 120</td>
<td>12 joints * 4 = 48</td>
<td>168</td>
</tr>
<tr>
<td>mTSS</td>
<td>280</td>
<td>168</td>
<td>448</td>
</tr>
</tbody>
</table>

2) Handling of Missing Joint Data

This section is focused on how to handle missing scores for individual joints during mTSS derivation.

If repeated X-ray was taken at screening, Week 12, Week 24 or early termination visit, the joint scores from the repeated X-ray will be set to missing if (repeated X-ray date - original X-ray date) > 28 days. At Week 52 the joint scores from the repeat X-ray will be set to missing if (repeated X-ray date - original X-ray date) > 42 days. If a joint was identified as surgically modified at screening, this joint will be set to be missing at all following visits.

Step 1 is to identify missing joint scores and impute values when applicable for each reader. The joints are divided into 8 segments. A segment is adequate when the total number of joints with nonmissing score is above the threshold as defined in the table below.

The change score of missing joints within an segment will be imputed based on the mean change scores of the available joints of that segment. After the missing joint change scores are imputed, the sum of joint change scores will be computed for each segment.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Total number of joints</th>
<th>Adequacy threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIP Erosion</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>MCP and Thumb Erosion</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Wrist Erosion</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Foot Erosion</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>PIP JSN</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>MCP and Thumb JSN</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Wrist JSN</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Foot JSN</td>
<td>12</td>
<td>7</td>
</tr>
</tbody>
</table>

Step 2 is to identify missing timepoint and impute values when applicable for each reader.
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis Version 1.0

The mTSS at baseline will be set to missing if any of the segments is inadequate. The mTSS at postbaseline visits will be set to missing if more than 3 segments are inadequate; otherwise imputation will be applied to imputed segments. The change score of the missing segments at a timepoint will be imputed based on the mean change scores of the available segments of that timepoint.

The erosion score and JSN score at baseline will be set to missing if any of the segments that contribute to derive either the erosion score or JSN score is inadequate, respectively. The erosion score and JSN score at postbaseline visits will be set to missing if more than 1 segment is inadequate, respectively; otherwise the change score of the missing segment will be imputed in the similar way as mTSS.

The missing joint scores at baseline will be imputed when all of the segments at baseline are adequate and the corresponding joints are scored at Week 24. If the corresponding segment at Week 24 is also adequate, the missing joint change scores at Week 24 will be imputed first using the method described before, and the missing joint score at baseline will be imputed as joint score at Week 24 minus the change in joint score. The mTSS at baseline will be the sum of the scores after imputation. The mTSS at postbaseline visits will be derived as change in mTSS plus baseline mTSS. The erosion score and JSN score at baseline and postbaseline visits will be derived in the similar way as mTSS.

Step 3 is to identify which reader pair will be used to derive final mTSS, erosion score and JSN score when an adjudicated read is present. The selection is made at Week 24 based on changes in mTSS as specified below and will be applied to all the visits.

- If both readers and adjudicator have change in mTSS available, the selected reader will be the one whose change in mTSS is closest to that of adjudicator. If the differences of change in mTSS comparing to adjudicator are the same for both readers, then both readers will be selected.

- If adjudicator’s change in mTSS is available, but one of the readers’ change in mTSS is missing, then the other reader will be selected.

Step 4 is to derive the final mTSS and final change in mTSS at a given timepoint. In the case of not adjudicated subjects the final mTSS will be the mean of the mTSS generated by the 2 readers. For the final change in mTSS, the change from baseline in mTSS will first be calculated for the 2 readers separately, and the final change score will be the mean of these two change scores. In the case of adjudicated subjects the final mTSS and final change in mTSS will be the mean of the mTSS and change in mTSS generated by the selected reader(s) at step 3 and the adjudicator respectively. The final erosion score, JSN score, and change in erosion and JSN score will be derived in the similar way as mTSS.
Appendix 4. Treatment Satisfaction Questionnaire for Medication (TSQM)

TSQM scale scores are computed by adding the items loading on each factor. The lowest possible score is subtracted from this composite score and divided by the greatest possible score minus the lowest possible score. This provides a transformed score between 0 and 1 that should be multiplied by 100. Note that only one item may be missing from each scale before the subscale should be considered invalid for that respondent.

**EFFECTIVENESS:**

\[
\text{If one item is missing} \\
\frac{([\text{Sum(}1 + \text{Item 2} + \text{Item 3}) - 3])}{18} \times 100
\]

**SIDE EFFECTS:**

If Question 4 is answered: No, then score = 100

Else,

\[
\frac{([\text{Sum(}5 \text{ to Item 8}) - 4])}{16} \times 100
\]

**CONVENIENCE:**

\[
\text{If one item is missing} \\
\frac{([\text{Sum(}9 \text{ to Item 11}) - 3])}{18} \times 100
\]

**GLOBAL SATISFACTION:**

\[
\text{If either Item 12 or 13 is missing} \\
\frac{([\text{Sum(}12 \text{ to Item 14}) - 3])}{14} \times 100
\]

\[
\text{If Item 14 is missing} \\
\frac{([\text{Sum(}12 \text{ and Item 13}) - 2])}{8} \times 100
\]
Appendix 5. Corticosteroids

The following table will be used for converting non-prednisone medications to prednisone equivalent:

Example: Subject is taking 8 mg of Methylprednisolone orally daily. To get the equivalent dose of prednisone: 8 mg Methylprednisolone = (5*8)/4 = 10 mg prednisone.

<table>
<thead>
<tr>
<th>Corticosteroids Name</th>
<th>Equivalent Dose (mg) to 5 mg Prednisone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betamethasone</td>
<td>0.75</td>
</tr>
<tr>
<td>Betamethasone Dipropionate</td>
<td>0.75</td>
</tr>
<tr>
<td>Betamethasone Sodium Phosphate</td>
<td>0.75</td>
</tr>
<tr>
<td>Cortisone</td>
<td>20</td>
</tr>
<tr>
<td>Dexamethasone</td>
<td>0.75</td>
</tr>
<tr>
<td>Dexamethasone Palmitate</td>
<td>0.75</td>
</tr>
<tr>
<td>Dexamethasone Phosphate</td>
<td>0.75</td>
</tr>
<tr>
<td>Dexamethasone Sodium Phosphate</td>
<td>0.75</td>
</tr>
<tr>
<td>Hydrocortisone</td>
<td>20</td>
</tr>
<tr>
<td>Meprednisone</td>
<td>4</td>
</tr>
<tr>
<td>Methylprednisolone</td>
<td>4</td>
</tr>
<tr>
<td>Methylprednisolone Acetate</td>
<td>4</td>
</tr>
<tr>
<td>Methylprednisolone Sodium Succinate</td>
<td>4</td>
</tr>
<tr>
<td>Prednisone</td>
<td>5</td>
</tr>
<tr>
<td>Prednisolone</td>
<td>5</td>
</tr>
<tr>
<td>Prednisolone Farnesylate</td>
<td>5</td>
</tr>
<tr>
<td>Prednisolone Sodium Succinate</td>
<td>5</td>
</tr>
<tr>
<td>Triamcinolone</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix 6. SAS Programming for Tipping Point Analysis for Binary Endpoint

The following %tp_binary macro generates multiple imputed data and a set of the shift parameters that adjust the imputed values will be examined.

/*------------------------------------------------------------------------------------------*/
/*--- Delta-Adjusting Method for Tipping Point Analysis for Binary Endpoint */
/*--- Generate imputed data set for specified shift parameters */
/*--- data= input data set */
/*--- smin= min shift parameter for active drug */
/*--- smax= max shift parameter for active drug */
/*--- sinc= increment of the shift parameter for active drug */
/*--- pmn= min shift parameter for placebo drug */
/*--- pmx= max shift parameter for placebo drug */
/*--- pinc= increment of the shift parameter for placebo drug */
/*--- trt= treatment group indicator */
/*--- out= output imputed data set */
/*------------------------------------------------------------------------------------------*/

%macro tp_binary( data=, smin=, smax=, sinc=, pmn=, pmx=, pinc=, trt=, out=);
   data &out;
      set _null_;
      run;

   /*-------- # of shift values --------*/
   %let ncase_pbo = %sysevalf((&pmx-&pmn)/&pinc, ceil);

   %do pc=0 %to &ncase_pbo;
     %let pj= %sysevalf( &pmn + &pc * &pinc);

     /*-------- # of shift values --------*/
     %let ncase= %sysevalf((&smax-&smin)/&sinc, ceil);

     /*-------- Imputed data for each shift --------*/
     %do jc=0 %to &ncase;
       %let sj= %sysevalf( &smin + &jc * &sinc);

         proc mi data=&data seed=14823 nimitate=20 out=outmi;
            var trt01pn strat1V strat2V strat3V ACR20;
            class trt01pn strat1V strat2V strat3V ACR20;
            monotone logistic (ACR20 / link=glogit);
            rmvar adjust(ACR20 (event=1) / adjustobs=(trt01pn="&trt") shift= &sj) adjust(ACR20 (event=1) / adjustobs=(trt01pn="3") shift= &pj));
            run;

         data outmi;
            set outmi;
            Shift_Trt= &sj;
            Shift_Pbo= &pj;
            run;

         data &out;

%end;
%end;
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis  Version 1.0

set & out outmi;
run;
%end;
%end;
%mend tp_binary;
Appendix 7. SAS Programming for Tipping Point Analysis for Continuous Endpoint

The following %tp_conti macro generates multiple imputed data sets and a set of the shift parameters that adjust the imputed values will be examined.

```sas
%macro tp_conti(data=, smin=, smax=, sinc=, pmin=, pmax=, pinc=, trt=, out=);
  data &data;
  set _null_; run;
  %let ncase_pbo = %sysvalf((&pmax-&pmin)/&pinc, ceil);
  %do pc=0 %to &ncase_pbo;
    %let pj = %sysvalf(&pmin + &pc * &pinc);
    %let s1 = %sysvalf(&smin + &pc * &sinc);
    /********** # of shift values **********/
    %let ncase = %sysvalf((&smax-&smin)/&sinc, ceil);
    %do j=0 %to &ncase;
      %let s2 = %sysvalf(&smin + &j * &sinc);
      proc mi data=&data out=all_mono ninput=20 seed=123;
        var v_1 v_2 v_3 v_4 v_5 v_6 v_7 v_8 v_9;
        mcmc chain=multiple impute=monotone;
      run;
      proc sort data=all_mono by _Imputation_ trt01pn;run;
      proc mi data=allMono out=outmi ninput=465 ninput=1;
        by _Imputation_; var trt01pn strat1V strat2V strat3V v_1 v_2 v_3 v_4 v_5 v_6 v_7 v_8 v_9;
        class trt01pn strat1V strat2V strat3V;
        monotone regression;
        mivar adjust v_5 / adjustobs=(trt01pn="&trt") shift=&s1
        adjust(v_5 / adjustobs=(trt01pn="3") shift= &pj);```
```
Filgotinib
GS-US-417-0301
Statistical Analysis Plan - Week 24 Analysis and Final Analysis
Version 1.0

run;
data outmi;
   set outmi;
   Shift_Trt= &sj;
   Shift_Pbo= &pj;
run;
data &out;
   set &out outmi;
run;
%end;
%end tp_cont;

Appendix 8. Sample SAS Code for Analysis of Binary Endpoints

The following model statement will be used to construct the confidence interval for the binomial proportions described in Sections 6.2.3, 6.3.2 and 6.4.2:

```sas
proc freq data = test;
  by visit;
  tables TRT01PN*ava1 / riskdiff (CORRECT);
  output out = f200vp_ci RISKDIFF;
  where TRT01PN = 1 | TRT01PN = 3;
run;
```

where the confidence interval for the difference of two independent binomial proportions is constructed based on the normal approximation (i.e., the Wald method) with continuity correction, to adjust for the difference between the normal approximation and the binomial distribution, which is a discrete distribution:

\[
(\hat{p}_1 - \hat{p}_2) \pm \frac{1}{2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right) + z_{\alpha/2} \times \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}
\]

and the confidence interval for the binomial proportions is constructed based on the normal approximation to the binomial distribution with continuity correction:

\[
\hat{p}_1 \pm \frac{1}{2n_1} + z_{\alpha/2} \times \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1}}
\]
**STATISTICAL ANALYSIS PLAN ADDENDUM**

<table>
<thead>
<tr>
<th>Study Title:</th>
<th>A Randomized, Double-blind, Placebo- and Active-controlled, Multicenter, Phase 3 Study to Assess the Efficacy and Safety of Filgotinib Administered for 52 weeks in Combination with Methotrexate to Subjects with Moderately to Severely Active Rheumatoid Arthritis Who Have an Inadequate Response to Methotrexate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Test Drug:</td>
<td>Filgotinib</td>
</tr>
<tr>
<td>Study Number:</td>
<td>GS-US-417-0301</td>
</tr>
<tr>
<td>Protocol Version (Date):</td>
<td>Amendment 1 (05 July 2016)</td>
</tr>
<tr>
<td>Analysis Type:</td>
<td>Final Analysis</td>
</tr>
<tr>
<td>Analysis Plan Version:</td>
<td>Version 2.0</td>
</tr>
<tr>
<td>Analysis Plan Date:</td>
<td>12 August 2019</td>
</tr>
<tr>
<td>Analysis Plan Author(s):</td>
<td>Lei Ye, Ph.D.</td>
</tr>
</tbody>
</table>

**CONFIDENTIAL AND PROPRIETARY INFORMATION**
TABLE OF CONTENTS

TABLE OF CONTENTS ........................................................................................................2
LIST OF IN-TEXT TABLES .................................................................................................2
1. INTRODUCTION ...........................................................................................................3
2. EFFICACY ANALYSES ...............................................................................................4
   2.1. General Considerations .........................................................................................4
   2.2. Additional Analyses of Efficacy Endpoints per FDA Comments .........................8
   2.3. Additional Analyses of Low Disease Activity and Remission .............................9
3. ADDITIONAL SAFETY ANALYSES ...........................................................................10
4. SAP REVISION ............................................................................................................11
5. REFERENCE ...............................................................................................................12
6. APPENDICES .............................................................................................................13
   Appendix 1. Tables, Figures, and Listings (TFLs) for the Additional Analyses.............13

LIST OF IN-TEXT TABLES

Table 2-1. Estimands for Binary Endpoints: Filgotinib versus Placebo ................................5
Table 2-2. Estimands for Binary Endpoints: Filgotinib versus Adalimumab .....................6
Table 2-3. Estimands for Continuous Endpoints: Filgotinib versus Placebo ..................7
Table 2-4. Analysis Visit Windows for On-treatment mTSS Data ....................................9
Table 2-5. Analysis Visit Windows for On-treatment SF-36 and FACIT-Fatigue .............9
1. INTRODUCTION

This statistical analysis plan (SAP) addendum provides the additional efficacy analyses per the Food and Drug Administration (FDA) comments on the SAPs of Studies GS-US-417-0301 and GS-US-417-0303 dated 15 February 2019 (Reference ID: 4391728).

In addition, the following additional analysis through Week 52 will also be provided:

- Low disease activity (LDA) and Remission per Clinical Disease Activity Index (CDAI) and Simplified Disease Activity Index (SDAI) criteria
- Boolean remission
- Lymphocyte subset counts (TBNK cells)
- Immunoglobulins (Ig) A, G, M, and total Ig level

The detailed statistical methods and data presentations are defined in the original SAP of this study.
2. EFFICACY ANALYSES

2.1. General Considerations

Estimands:

Three efficacy estimands, composite estimand, treatment-policy estimand and hypothetical estimand are defined for the primary and key secondary efficacy endpoints, respectively, as described in Table 2-1, Table 2-2 and Table 2-3 below based on ICH E9R (1) (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2017).
Table 2-1. Estimands for Binary Endpoints: Filgotinib versus Placebo

<table>
<thead>
<tr>
<th>Binary Endpoint</th>
<th>Composite</th>
<th>Treatment Policy</th>
<th>Hypothetical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
</tr>
<tr>
<td>Patient-level outcome to be measured</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., A CR20 response at Week 12)</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., A CR20 response at Week 12)</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., A CR20 response at Week 12)</td>
</tr>
<tr>
<td>Measure of intervention effect &amp; handling of intercurrent events</td>
<td>Measure of the effect of the initially randomized treatments assuming that any rescue efforts (i.e., escape to standard of care) or premature discontinuations from treatment result in a non-response</td>
<td>Measure of the treatment effect regardless of what treatment was actually received and ignoring the occurrence of intercurrent events (i.e., protocol violations, escape to standard of care, or premature study discontinuation)</td>
<td>Measure of the effect of the initially randomized treatments assuming all patients had remained on their randomized treatment throughout the study (i.e., assuming patients did not receive rescue medication or did not discontinue study)</td>
</tr>
<tr>
<td>Population-level summary measure</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to placebo</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to placebo</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to placebo</td>
</tr>
<tr>
<td>Estimators</td>
<td><strong>Main Estimator</strong>: The treatment difference for a binary efficacy endpoint will be estimated using a logistic regression model. Missing data will be imputed using non-responder imputation.</td>
<td><strong>Main Estimator</strong>: The treatment difference for a binary efficacy endpoint will be estimated using a logistic regression model. Missing data will be imputed using multiple imputation. <strong>Sensitivity Estimator</strong>: The tipping point method will be used. <strong>Supportive Estimator</strong>: Only observed data will be used for this model. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study. No imputation will be performed.</td>
<td><strong>Main Estimator</strong>: The treatment difference for a binary efficacy endpoint will be estimated using a logistic regression model. Missing data will be imputed using multiple imputation (see bullet points below for additional details regarding missing data). If a subject takes standard of care medications due to inadequate response or discontinues from study treatment, the data will be treated as missing from the point of inadequate response or study drug discontinuation. If a subject discontinues from the study, treat data as missing from the point of loss-to-follow-up onward to the end of the study. <strong>Sensitivity Estimator</strong>: The tipping point method will be used. <strong>Supportive Estimator</strong>: Only observed data will be used for this model. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study.</td>
</tr>
<tr>
<td>Binary Endpoint</td>
<td>Composite</td>
<td>Treatment Policy</td>
<td>Hypothetical</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>-----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Population</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
</tr>
<tr>
<td>Patient-level outcome to be measured</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., DAS28(CRP) ( \geq 3.2 ); DAS28(CRP) &lt; 2.6)</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., DAS28(CRP) ( \geq 3.2 ); DAS28(CRP) &lt; 2.6)</td>
<td>Outcome Response for a binary endpoint at Week 12 (e.g., DAS28(CRP) ( \geq 3.2 ); DAS28(CRP) &lt; 2.6)</td>
</tr>
<tr>
<td>Measure of the intervention effect &amp; handling of intercurrent events</td>
<td>Measure of the effect of the initially randomized treatments assuming that any rescue efforts (i.e., escape to standard of care) or premature discontinuations from treatment result in a non-response</td>
<td>Measure of the treatment effect regardless of what treatment was actually received and ignoring the occurrence of intercurrent events (i.e., protocol violations, escape to standard of care, or premature study discontinuation)</td>
<td>Measure of the effect of the initially randomized treatments assuming all patients had remained on their randomized treatment throughout the study (i.e., assuming patients did not receive rescue medication or did not discontinue study)</td>
</tr>
<tr>
<td>Population-level summary measure</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to adalimumab</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to adalimumab</td>
<td>Difference in a binary efficacy endpoint, comparing those assigned to filgotinib to those assigned to adalimumab</td>
</tr>
</tbody>
</table>

**Estimators**
- **Main Estimator**: The non-inferiority of filgotinib over adalimumab for a binary efficacy endpoint will be estimated using non-inferiority test. Missing data will be imputed using non-responder imputation.
- **Sensitivity Estimator**: The tipping point method will be used.
- **Supportive Estimator**: Only observed data will be used for this model. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study. No imputation will be performed.
### Table 2-3. Estimands for Continuous Endpoints: Filgotinib versus Placebo

<table>
<thead>
<tr>
<th>Continuous Endpoint</th>
<th>Composite</th>
<th>Treatment Policy</th>
<th>Hypothetical</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
<td>Full Analysis Set as defined in the protocols (Sections 4.2, 4.3, and 8.2.1)</td>
</tr>
<tr>
<td><strong>Patient-level outcome to be measured</strong></td>
<td>Change from baseline in the continuous outcome at Week 12 or Week 24 (e.g., Change from baseline in mTSS at Week 24)</td>
<td>Change from baseline in the continuous outcome at Week 12 or Week 24 (e.g., Change from baseline in mTSS at Week 24)</td>
<td>Change from baseline in the continuous outcome at Week 12 or Week 24 (e.g., Change from baseline in mTSS at Week 24)</td>
</tr>
<tr>
<td><strong>Measure of intervention effect &amp; handling of intercurrent events</strong></td>
<td>Measure of the effect of the initially randomized treatments assuming that any rescue efforts (i.e., escape to standard of care) or premature discontinuations from treatment result in an unfavorable outcome</td>
<td>Measure of the treatment effect regardless of what treatment was actually received and ignoring the occurrence of intercurrent events (i.e., protocol violations, escape to standard of care, or premature study discontinuation)</td>
<td>Measure of the effect of the initially randomized treatments assuming all patients had remained on their randomized treatment throughout the study (i.e., assuming patients did not receive rescue medication or did not discontinue study)</td>
</tr>
<tr>
<td><strong>Population-level summary measure</strong></td>
<td>Mean difference between treatment arms (comparing those assigned to filgotinib to those assigned to placebo) in the change from baseline to Week 12 or Week 24 in a given continuous outcome</td>
<td>Mean difference between treatment arms (comparing those assigned to filgotinib to those assigned to placebo) in the change from baseline to Week 12 or Week 24 in a given continuous outcome</td>
<td>Mean difference between treatment arms (comparing those assigned to filgotinib to those assigned to placebo) in the change from baseline to Week 12 or Week 24 in a given continuous outcome</td>
</tr>
<tr>
<td><strong>Estimator(s)</strong></td>
<td>Main Estimator: The treatment difference for a continuous efficacy endpoint will be estimated using a mixed effects model. Missing data will be imputed using multiple imputation based on observed data from the placebo group.</td>
<td>Main Estimator: The treatment difference for a continuous efficacy endpoint will be estimated using a mixed effects model. Missing data will be imputed using multiple imputation. Sensitivity Estimator: The tipping point method will be used. Supportive Estimator: An ANCOVA model using observed data only will be fit. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study. No imputation will be performed.</td>
<td>Main Estimator: The treatment difference for a continuous efficacy endpoint will be estimated using a mixed effects model. Missing data will be imputed using multiple imputation (see bullet points below for additional details regarding missing data). If a subject takes standard of care medications due to inadequate response or discontinues from study treatment, the data will be treated as missing from the point of inadequate response or study drug discontinuation. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study. Sensitivity Estimator: The tipping point method will be used.</td>
</tr>
</tbody>
</table>
2.2. Additional Analyses of Efficacy Endpoints per FDA Comments

The following additional analyses will be provided for estimands:

- For composite estimand for continuous endpoints, the main estimator will be performed using the multiple imputation approach with missing data being imputed based on data observed from the placebo group.
  
  — For key continuous endpoints (change from baseline in HAQ-DI, SF-36 PCS, FACIT-F at Week 12 and mTSS at Week 24), multiple imputation will be performed by generating the multiple imputed datasets based on linear regression models on observed scores from the placebo group. These multiple imputed datasets are then analyzed by using the analysis method specified in the original SAP Section 6.3.2 for complete data. The results from each set of imputed datasets will then be combined using Rubin’s rule (Rubin 1987).

- For treatment policy estimand, the supportive estimator will be performed based on an ANCOVA model using only observed data:
  
  — For key continuous efficacy endpoints (change from baseline in HAQ-DI, SF-36 PCS, FACIT-F at Week 12 and mTSS at Week 24), an ANCOVA model with baseline value, stratification factors and treatment group in the model will be used for comparing a treatment difference between each filgotinib group and placebo in change from baseline to the time point of interest. The LS means along with the two-sided 95% CIs and the p-value for the difference in mean change from Baseline at the time point of interest from the model will be presented. The observed values will be used for analysis. If a subject discontinues from the study, the data will be treated as missing from the point of loss-to-follow-up to the end of the study. No imputation will be performed.

The following supportive analyses will be provided based on the FDA comments:

- Per the FDA comment that the MMRM model should also include subjects with baseline measurement available, a revised MMRM model will be performed for key continuous efficacy endpoints (change from baseline in HAQ-DI, SF-36 PCS, FACIT-F at Week 12 and mTSS at Week 24). The model will include all the subject with baseline measurement with stratification factors, treatment, visit, and treatment by visit interaction included as fixed effects and subject as the random effect. The model will use actual observations including baseline as outcome and the baseline will not be adjusted as covariate. This analysis will be based on the on-treatment data and all available data as described in Section 6.1 of the original SAP.

- Per the FDA comment on the analysis window, the following changes will be made on the window definition: for mTSS and PRO endpoints (SF-36 PCS, and FACIT-F), visit windows will include +/- 4 weeks as described in Table 2-4 and Table 2-5. The analysis will follow the same analysis method specified for primary analysis of these endpoints as described in Section 6.3.2 of the original SAP.
Table 2-4. Analysis Visit Windows for On-treatment mTSS Data

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>15</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>58</td>
<td>113</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>142</td>
<td>197</td>
</tr>
</tbody>
</table>

Table 2-5. Analysis Visit Windows for On-treatment SF-36 and FACIT-Fatigue

<table>
<thead>
<tr>
<th>Analysis Visit</th>
<th>Nominal Study Day</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>(none)</td>
<td>1</td>
</tr>
<tr>
<td>Week 4</td>
<td>29</td>
<td>2</td>
<td>57</td>
</tr>
<tr>
<td>Week 12</td>
<td>85</td>
<td>58</td>
<td>113</td>
</tr>
<tr>
<td>Week 24</td>
<td>169</td>
<td>142</td>
<td>197</td>
</tr>
</tbody>
</table>

2.3. Additional Analyses of Low Disease Activity and Remission

The following additional efficacy endpoints will be analyzed:

- Low disease activity (LDA) per CDAI or SDAI criteria is defined as CDAI \( \geq 10 \) or SDAI \( \geq 11 \).

- Remission per CDAI or SDAI criteria is defined as CDAI \( \geq 2.8 \) or SDAI \( \geq 3.3 \).

- Boolean remission is achieved when tender joint count 28 (TJC28) \( \geq 1 \) and swollen joint count 28 (SJC28) \( \geq 1 \) and c-reactive protein (CRP) \( \geq 1 \) mg/dL and subject's global assessment (SGA) \( \geq 1 \) (on a 0-10 cm scale). Boolean remission is not achieved when at least 1 of the 4 components has a value \( > 1 \). In case that some components are missing, Boolean remission is not achieved if at least 1 of the non-missing components has a value \( > 1 \). If non-missing components are not sufficient to determine Boolean remission, then Boolean remission will be considered as missing.

The proportion of subjects who achieved LDA or Remission per the above criteria will be analyzed using main estimator of treatment policy estimand and composite estimand, respectively. In addition, the proportion of subjects who achieved LDA or Remission will be analyzed through Week 24 using the logistic regression method with NRI.
3. ADDITIONAL SAFETY ANALYSES

The descriptive statistics of Baseline values, values at each postbaseline visit and change from Baseline at each postbaseline visit will be provided by treatment groups for the following laboratory tests.

- Lymphocyte subset counts (TBNK cells)
  - CD3 (%)
  - CD3 (/μL)
  - CD4 (%)
  - CD4 (/μL)
  - CD8 (%)
  - CD8 (/μL)
  - CD19 (%)
  - CD19 (/μL)
  - CD16+56 (%)
  - CD16+56 (/μL)

- Immunoglobulins (Ig) A, G, M, and total Ig level

In addition, the proportion of subjects below the lower limit of normal (LLN) or above the upper limit of normal (ULN) will be summarized descriptively for TBNK cells. The proportion of subjects below the LLN will be summarized descriptively for Immunoglobulins (Ig) A, G, and M.

The list of tables, figures, and listings (TFLs) for the additional analyses is shown in Appendix 1.
### 4. SAP REVISION

<table>
<thead>
<tr>
<th>Revision Date (DD MMM YYYY)</th>
<th>Section</th>
<th>Summary of Revision</th>
<th>Reason for Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 August 2019</td>
<td>Section 2.1</td>
<td>Removal of treatment policy main estimator as the primary analysis and added as supportive analysis</td>
<td>Based on the Division’s feedback at the 07 August 2019 Type C meeting, the prespecified analysis will remain as the primary analysis.</td>
</tr>
</tbody>
</table>
5. REFERENCE


# APPENDICES

## Appendix 1. Tables, Figures, and Listings (TFLs) for the Additional Analyses

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HAQ-DI: Change from Baseline at Week 12 - Multiple Imputation with Unfavorable Outcome</td>
</tr>
<tr>
<td>2</td>
<td>HAQ-DI: Change from Baseline at Week 12 - All Available Data Included (ANCOVA analysis)</td>
</tr>
<tr>
<td>3</td>
<td>HAQ-DI: Change from Baseline at Week 12 - MMRM including Subjects with Baseline Value Available</td>
</tr>
<tr>
<td>4</td>
<td>HAQ-DI: Change from Baseline at Week 12 - All Available Data Included (MMRM including Subjects with Baseline Value Available)</td>
</tr>
<tr>
<td>5</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - Multiple Imputation with Unfavorable Outcome</td>
</tr>
<tr>
<td>6</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - All Available Data Included (ANCOVA analysis)</td>
</tr>
<tr>
<td>7</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - MMRM including Subjects with Baseline Value Available</td>
</tr>
<tr>
<td>8</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - All Available Data Included (MMRM including Subjects with Baseline Value Available)</td>
</tr>
<tr>
<td>9</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - Multiple Imputation with Unfavorable Outcome</td>
</tr>
<tr>
<td>10</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - All Available Data Included (ANCOVA analysis)</td>
</tr>
<tr>
<td>11</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - MMRM including Subjects with Baseline Value Available</td>
</tr>
<tr>
<td>12</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - All Available Data Included (MMRM including Subjects with Baseline Value Available)</td>
</tr>
<tr>
<td>13</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - Multiple Imputation with Unfavorable Outcome</td>
</tr>
<tr>
<td>14</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - All Available Data Included (ANCOVA analysis)</td>
</tr>
<tr>
<td>15</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - MMRM including Subjects with Baseline Value Available</td>
</tr>
<tr>
<td>16</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - All Available Data Included (MMRM including Subjects with Baseline Value Available)</td>
</tr>
<tr>
<td>17</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - Revised Analysis Window</td>
</tr>
<tr>
<td>18</td>
<td>van der Heijde modified Total Sharp Score (mTSS): Change from Baseline at Week 24 - All Available Data Included (Revised Analysis Window)</td>
</tr>
<tr>
<td>19</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - Revised Analysis Window</td>
</tr>
<tr>
<td>20</td>
<td>SF-36 Physical Component Summary (PCS): Change from Baseline at Week 12 - All Available Data Included (Revised Analysis Window)</td>
</tr>
</tbody>
</table>
### Table

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - Revised Analysis Window</td>
</tr>
<tr>
<td>22</td>
<td>FACIT-Fatigue: Change from Baseline at Week 12 - All Available Data Included (Revised Analysis Window)</td>
</tr>
<tr>
<td>23</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 10 by Visit through Week 24: Pairwise Comparisons versus Placebo (Logistic Regression with NRI)</td>
</tr>
<tr>
<td>24</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 10 by Visit from Week 24 to Week 52 - All Available Data Included</td>
</tr>
<tr>
<td>25</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 2.8 by Visit through Week 24: Pairwise Comparisons versus Placebo (Logistic Regression with NRI)</td>
</tr>
<tr>
<td>26</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 2.8 by Visit from Week 24 to Week 52 - All Available Data Included</td>
</tr>
<tr>
<td>27</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 11 by Visit through Week 24: Pairwise Comparisons versus Placebo (Logistic Regression with NRI)</td>
</tr>
<tr>
<td>28</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 11 by Visit from Week 24 to Week 52 - All Available Data Included</td>
</tr>
<tr>
<td>29</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 3.3 by Visit through Week 24: Pairwise Comparisons versus Placebo (Logistic Regression with NRI)</td>
</tr>
<tr>
<td>30</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 3.3 by Visit from Week 24 to Week 52 - All Available Data Included</td>
</tr>
<tr>
<td>31</td>
<td>Boolean Remission by Visit through Week 24: Pairwise Comparisons versus Placebo (Logistic Regression with NRI)</td>
</tr>
<tr>
<td>32</td>
<td>Boolean Remission by Visit from Week 24 to Week 52 - All Available Data Included</td>
</tr>
<tr>
<td>33</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 2.8 at Week 12: Pairwise Comparisons versus Placebo - All Available Data Included (Logistic Regression with Multiple Imputation)</td>
</tr>
<tr>
<td>34</td>
<td>Clinical Disease Activity Index (CDAI) &lt;= 10 at Week 12: Pairwise Comparisons versus Placebo - All Available Data Included (Logistic Regression with Multiple Imputation)</td>
</tr>
<tr>
<td>35</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 3.3 at Week 12: Pairwise Comparisons versus Placebo - All Available Data Included (Logistic Regression with Multiple Imputation)</td>
</tr>
<tr>
<td>36</td>
<td>Simplified Disease Activity Index (SDAI) &lt;= 11 at Week 12: Pairwise Comparisons versus Placebo - All Available Data Included (Logistic Regression with Multiple Imputation)</td>
</tr>
<tr>
<td>37</td>
<td>Boolean Remission at Week 12: Pairwise Comparisons versus Placebo - All Available Data Included (Logistic Regression with Multiple Imputation)</td>
</tr>
<tr>
<td>38</td>
<td>Cellmark: TBNK cell parameter Laboratory Events by Visit</td>
</tr>
<tr>
<td>39</td>
<td>Cellmark: TBNK cell parameter and Change from Baseline</td>
</tr>
<tr>
<td>40</td>
<td>Chemistry: Immunoglobulin A (mg/dL) Laboratory Events by Visit</td>
</tr>
<tr>
<td>41</td>
<td>Chemistry: Immunoglobulin G (mg/dL) Laboratory Events by Visit</td>
</tr>
<tr>
<td>42</td>
<td>Chemistry: Immunoglobulin M (mg/dL) Laboratory Events by Visit</td>
</tr>
<tr>
<td>43</td>
<td>Chemistry: Immunoglobulin A (mg/dL) and Change from Baseline</td>
</tr>
<tr>
<td>44</td>
<td>Chemistry: Immunoglobulin G (mg/dL) and Change from Baseline</td>
</tr>
<tr>
<td>45</td>
<td>Chemistry: Immunoglobulin M (mg/dL) and Change from Baseline</td>
</tr>
<tr>
<td>46</td>
<td>Chemistry: Total Immunoglobulin (mg/dL) and Change from Baseline</td>
</tr>
</tbody>
</table>
SUPPLEMENTARY MATERIAL

Supplementary methods

Assessment of radiographic progression based on van der Heijde modified total Sharp score

The van der Heijde modified total Sharp score (mTSS) was assigned by 2 central readers blinded to treatment, patient information, and chronological order of the radiographs; average value is reported. Results reported for baseline, week 12, and week 24 were based on readings of radiographs obtained at screening, week 12, and week 24 or pretermination visit before week 24 (Campaign A). Week 52 results were based on readings of radiographs obtained at screening, week 24, and week 52 or pretermination visit after week 24 (Campaign B) combined with Campaign A results.

Statistical analysis

Noninferiority analyses

The noninferiority test of the proportions of patients achieving Disease Activity Score in 28 joints with C-reactive protein (DAS28[CRP]) ≤3.2 and <2.6 at week 12 assessed whether each filgotinib dose preserves more than 50% of the effect of adalimumab compared with placebo (minimum effect fraction = 0.5). The minimum effect fraction of 0.5 was chosen based on published methods suggesting a fraction of 0.5–0.99 for noninferiority and a similar trial of tofacitinib employing a noninferiority margin based on 50% of the treatment difference for adalimumab[1-3].

Sensitivity analyses

The multiple imputation procedure replaced each missing value with a set of plausible values that represented the uncertainty about the right value to impute. Fifty imputed data sets were generated based on logistic regression models for binary efficacy endpoints (eg, ACR20) or linear regression models for continuous efficacy end points (eg, Health Assessment Questionnaire-Disability Index). These multiple imputed data sets were analysed using the same method as for the primary analysis. The results from each set of imputed data sets were combined using Rubin’s rule.[4] The stratification factors were included in...
the imputation model as covariates, and all available data at postbaseline visits up to the time point of interest were included in the model.

All statistical analyses were done using SAS version 9.4 (SAS Institute).

Supplementary results

Narratives of deaths

1. The acute DVT-associated death occurred in a patient in their 60s with risk factors of obesity, history of tobacco use, African descent, and cardiovascular disease. The patient started placebo on study day 1 and was rerandomised to filgotinib 200 mg (FIL200) on study day 169. The patient received FIL200 for 37 days before study drug discontinuation on day 205 due to persistent anaemia, radiographic finding of pneumoperitoneum, and weight loss. Investigation did not identify an underlying cause. The patient experienced ischaemic stroke and DVT considered related to study drug on day 219 (13 days after the last dose of filgotinib), followed by pulmonary embolism on day 220, and death on day 224. Investigation revealed bilateral pulmonary emboli with scattered interstitial opacities of the bilateral upper and lower lobes concerning for pulmonary infarcts and deep venous thrombosis of the femoral vein (thrombus within the bilateral popliteal veins with extension into the inferior femoral vein on the left). There was no evidence of right heart strain. The patient was also found to have cardiomegaly with patent foramen ovale, which could explain the distal internal carotid artery, middle cerebral artery, and posterior cerebral artery occlusions.

2. The primary varicella infection-associated death occurred in a patient in their 30s with unknown history of childhood chicken pox or varicella vaccination who completed study treatment with filgotinib 100 mg (after rerandomisation from placebo) on day 361—the end of the study—and continued receiving methotrexate and corticosteroids for treatment of rheumatoid arthritis. The patient presented 5 days after study drug discontinuation with a disseminated vascular rash diagnosed as varicella based on appearance. It is unknown if the patient had ever received varicella vaccination; however, the investigator noted that vaccination for chicken pox was uncommon in that region and there were 2 other patients...
admitted to the same hospital with varicella infection within ~1 week of this event. The patient’s death on day 368 was attributed to varicella and was considered related to study drug. No virology testing or autopsy was performed.
Supplementary References


Table S1. Noninferiority margins based on the difference of proportions of patients achieving DAS28(CRP) ≤3.2 or <2.6 between the adalimumab and placebo groups at weeks 12 and 24

<table>
<thead>
<tr>
<th></th>
<th>ADA (%)</th>
<th>PBO (%)</th>
<th>Noninferiority margin(^a) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAS28(CRP) ≤3.2, week 12</td>
<td>43.4</td>
<td>23.4</td>
<td>10.0</td>
</tr>
<tr>
<td>DAS28(CRP) ≤3.2, week 24</td>
<td>50.5</td>
<td>33.7</td>
<td>8.4</td>
</tr>
<tr>
<td>DAS28(CRP) &lt;2.6, week 12</td>
<td>23.7</td>
<td>9.3</td>
<td>7.2</td>
</tr>
<tr>
<td>DAS28(CRP) &lt;2.6, week 24</td>
<td>35.7</td>
<td>16.2</td>
<td>9.8</td>
</tr>
</tbody>
</table>

\(^a\)50% of the difference between ADA vs PBO.

ADA, adalimumab; DAS28(CRP), Disease Activity Score in 28 joints with C-reactive protein; PBO, placebo.
Table S2. Multiple imputation sensitivity analyses: Primary and secondary outcomes

<table>
<thead>
<tr>
<th></th>
<th>FIL200 (n = 475)</th>
<th>FIL100 (n = 480)</th>
<th>ADA (n = 325)</th>
<th>PBO (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR20, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>79.1 (75.4 to 82.9)</td>
<td>73.9 (69.9 to 77.9)</td>
<td>73.8 (68.9 to 78.6)</td>
<td>53.3 (48.7 to 57.9)</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)</td>
<td>25.9 (20.0 to 31.8)</td>
<td>20.6 (14.5 to 26.7)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key secondary outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ-DI change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>462</td>
<td>460</td>
<td>315</td>
<td>443</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−0.69 ± 0.61</td>
<td>−0.56 ± 0.57</td>
<td>−0.61 ± 0.56</td>
<td>−0.42 ± 0.54</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)</td>
<td>−0.29 (−0.35 to −0.22)</td>
<td>−0.17 (−0.24 to −0.10)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28(CRP) &lt;2.6, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>34.5 (30.2 to 38.8)</td>
<td>25.6 (21.6 to 29.6)</td>
<td>25.0 (20.3 to 29.8)</td>
<td>9.8 (7.1 to 12.5)</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)</td>
<td>24.7 (19.7 to 29.8)</td>
<td>15.8 (11.0 to 20.7)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mTSS change from baseline to week 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>428</td>
<td>429</td>
<td>283</td>
<td>390</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>0.15 ± 1.1</td>
<td>0.17 ± 0.89</td>
<td>0.17 ± 0.94</td>
<td>0.36 ± 1.4</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)</td>
<td>−0.24 (−0.39 to −0.08)</td>
<td>−0.23 (−0.39 to −0.07)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>0.003</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noninferiority DAS28(CRP) ≤3.2, week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI)</td>
<td>50.6 (46.1 to 55.1)</td>
<td>41.5 (37.0 to 46.0)</td>
<td>45.4 (39.9 to 50.8)</td>
<td>24.3 (20.4 to 28.2)</td>
</tr>
<tr>
<td>Difference vs ADA (95% CI)</td>
<td>5.2 (−1.9 to 12.3)</td>
<td>−3.9 (−10.9 to 3.2)</td>
<td>&lt;0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 PCS change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>465</td>
<td>465</td>
<td>315</td>
<td>448</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>9.1 ± 8.1</td>
<td>8.4 ± 7.7</td>
<td>8.4 ± 7.8</td>
<td>5.7 ± 7.2</td>
</tr>
<tr>
<td>Difference vs PBO (95% CI)</td>
<td>3.7 (2.8 to 4.6)</td>
<td>3.1 (2.2 to 4.0)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FACIT-F change from baseline to week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Mean ± SD</td>
<td>Difference vs PBO (95% CI)</td>
<td>Exploratory p vs PBO</td>
<td>Superiority DAS28(CRP) &lt;3.2, week 12</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>----------------------------</td>
<td>----------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td></td>
<td>458</td>
<td>9.2 ± 9.8</td>
<td>3.0 (2.0 to 4.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>9.0 ± 10.2</td>
<td>2.8 (1.7 to 3.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>309</td>
<td>8.8 ± 9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>440</td>
<td>6.6 ± 9.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Difference shown as difference in response rates for categorical outcomes and least-squares means difference for continuous outcomes.

ACR20, 20% improvement from baseline in American College of Rheumatology core criteria; ADA, adalimumab; CI, confidence interval; DAS28(CRP), Disease Activity Score in 28 joints with C-reactive protein; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; HAQ-DI, Health Assessment Questionnaire-Disability Index; mTSS, van der Heijde modified total Sharp score; PBO, placebo; SD, standard deviation; SF-36 PCS, Short Form-36 Physical Component Summary.
Table S3. Additional efficacy data through week 52

<table>
<thead>
<tr>
<th></th>
<th>FIL200 (n = 475)</th>
<th>FIL100 (n = 480)</th>
<th>ADA (n = 325)</th>
<th>PBO (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACR20</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>78.1 (74.3–81.9)</td>
<td>77.7 (73.9–81.5)</td>
<td>74.5 (69.6–79.4)</td>
<td>59.2 (54.6–63.7)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>18.9 (13.0–24.9)</td>
<td>18.6 (12.6–24.5)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>78.3 (74.5–82.1)</td>
<td>75.6 (71.7–79.6)</td>
<td>73.5 (68.6–78.5)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>4.8 (~1.5 to 11.1)</td>
<td>2.1 (~4.3 to 8.5)</td>
<td>0.10</td>
<td>0.52</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACR50</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>57.9 (53.3–62.4)</td>
<td>52.7 (48.1–57.3)</td>
<td>52.3 (46.7–57.9)</td>
<td>33.3 (28.9–37.6)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>24.6 (18.3–31.0)</td>
<td>19.4 (13.1–25.8)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>62.3 (57.9–66.8)</td>
<td>58.5 (54.0–63.1)</td>
<td>59.1 (53.6–64.6)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>3.2 (~3.9 to 10.4)</td>
<td>~0.5 (~7.7 to 6.7)</td>
<td>0.32</td>
<td>0.87</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACR70</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>36.4 (32.0–40.9)</td>
<td>30.2 (26.0–34.4)</td>
<td>30.2 (25.0–35.3)</td>
<td>15.4 (12.0–18.7)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>21.1 (15.4–26.7)</td>
<td>14.8 (9.4–20.3)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>44.2 (39.6–48.8)</td>
<td>37.5 (33.1–41.9)</td>
<td>39.4 (33.9–44.9)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>4.8 (~2.4 to 12.0)</td>
<td>~1.9 (~9.0 to 5.2)</td>
<td>0.15</td>
<td>0.61</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HAQ-DI change from baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>418</td>
<td>423</td>
<td>283</td>
<td>376</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−0.82 ± 0.63</td>
<td>−0.75 ± 0.60</td>
<td>−0.78 ± 0.63</td>
<td>−0.62 ± 0.60</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.27 (~−0.34 to ~−0.19)</td>
<td>−0.19 (~−0.26 to ~−0.11)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>400</td>
<td>398</td>
<td>265</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−0.92 ± 0.65</td>
<td>−0.85 ± 0.62</td>
<td>−0.85 ± 0.65</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0.09 (~−0.17 to ~−0.01)</td>
<td>−0.01 (~−0.09 to ~0.08)</td>
<td>0.036</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>DAS28(CRP) ≤2.6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>% (95% CI), week 24</strong></td>
<td>48.4 (43.8–53.0)</td>
<td>35.2 (30.8–39.6)</td>
<td>35.7 (30.3–41.1)</td>
<td>16.2 (12.8–19.6)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>32.2 (26.4–38.0)</td>
<td>19.0 (13.4–24.6)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>% (95% CI), week 52</strong></td>
<td>53.9 (49.3–58.5)</td>
<td>42.9 (38.4–47.4)</td>
<td>46.2 (40.6–51.7)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>7.7 (0.4–15.0)</td>
<td>3.2 (–10.5 to 4.0)</td>
<td>0.024</td>
<td>0.38</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>DAS28(CRP) ≤3.2</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>% (95% CI), week 24</strong></td>
<td>60.6 (56.1–65.1)</td>
<td>53.1 (48.6–57.7)</td>
<td>50.5 (44.9–56.1)</td>
<td>33.7 (29.3–38.0)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>26.9 (20.6–33.3)</td>
<td>19.4 (13.1–25.8)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>% (95% CI), week 52</strong></td>
<td>65.9 (61.5–70.3)</td>
<td>59.4 (54.9–63.9)</td>
<td>58.8 (53.3–64.3)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>7.1 (0–14.2)</td>
<td>0.6 (–6.6 to 7.8)</td>
<td>0.031</td>
<td>0.86</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>SF-36 PCS change from baseline</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>424</td>
<td>426</td>
<td>283</td>
<td>376</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>10.4 ± 8.5</td>
<td>10.3 ± 8.6</td>
<td>10.4 ± 8.5</td>
<td>7.7 ± 8.0</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>3.1 (2.1–4.1)</td>
<td>3.1 (2.0–4.1)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>399</td>
<td>398</td>
<td>265</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>12.0 ± 8.7</td>
<td>11.5 ± 8.7</td>
<td>12.4 ± 9.2</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>0.1 (–1.1 to 1.3)</td>
<td>–0.3 (–1.5 to 0.9)</td>
<td>0.84</td>
<td>0.64</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>FACIT-F change from baseline</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>413</td>
<td>417</td>
<td>273</td>
<td>369</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>10.5 ± 10.6</td>
<td>10.8 ± 10.8</td>
<td>10.3 ± 9.7</td>
<td>8.4 ± 10.5</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>2.6 (1.5–3.8)</td>
<td>2.8 (1.6–3.9)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>384</td>
<td>376</td>
<td>254</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>11.9 ± 10.2</td>
<td>12.2 ± 10.9</td>
<td>11.7 ± 10.8</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>0.7 (–0.6 to 2.0)</td>
<td>0.8 (–0.5 to 2.1)</td>
<td>0.28</td>
<td>0.22</td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDAI ≤10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>% (95% CI), week 12</td>
<td>45.9 (41.3–50.5)</td>
<td>36.7 (32.3–41.1)</td>
<td>38.5 (33.0–43.9)</td>
<td>23.8 (19.9–27.7)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>22.1 (16.0–28.2)</td>
<td>12.9 (6.9–18.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p for noninferiority vs ADA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>60.2 (55.7–64.7)</td>
<td>50.8 (46.3–55.4)</td>
<td>49.5 (43.9–55.1)</td>
<td>35.6 (31.2–40.0)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>24.6 (18.3–31.0)</td>
<td>15.3 (8.8–21.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p for noninferiority vs ADA</td>
<td>&lt;0.001</td>
<td>0.045</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>66.9 (62.6–71.3)</td>
<td>59.2 (54.7–63.7)</td>
<td>61.2 (55.8–66.7)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>5.7 (−1.3 to 12.8)</td>
<td>−2.1 (−9.2 to 5.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.077</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDAI ≤11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 12</td>
<td>46.9 (42.4–51.5)</td>
<td>36.7 (32.3–41.1)</td>
<td>38.5 (33.0–43.9)</td>
<td>24.4 (20.5–28.4)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>22.5 (16.4–28.7)</td>
<td>12.2 (6.2–18.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>60.2 (55.7–64.7)</td>
<td>52.5 (47.9–57.1)</td>
<td>50.5 (44.9–56.1)</td>
<td>35.4 (31.0–39.8)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>24.8 (18.5–31.2)</td>
<td>17.1 (10.7–23.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>67.4 (63.0–71.7)</td>
<td>59.6 (55.1–64.1)</td>
<td>60.0 (54.5–65.5)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>7.4 (0.3–14.4)</td>
<td>−0.4 (−7.6 to 6.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.025</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDAI ≤2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 12</td>
<td>12.4 (9.3–15.5)</td>
<td>11.0 (8.1–13.9)</td>
<td>5.8 (3.1–8.6)</td>
<td>2.7 (1.2–4.3)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>9.7 (6.2–13.2)</td>
<td>8.3 (4.9–11.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p for noninferiority vs ADA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>21.3 (17.5–25.0)</td>
<td>18.5 (15.0–22.1)</td>
<td>16.9 (12.7–21.2)</td>
<td>8.4 (5.8–11.0)</td>
</tr>
<tr>
<td>Treatment difference vs PBO (95% CI)</td>
<td>12.8 (8.2–17.5)</td>
<td>10.1 (5.6–14.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p for noninferiority vs ADA</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>29.5 (25.3–33.7)</td>
<td>24.2 (20.2–28.1)</td>
<td>22.8 (18.1–27.5)</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA (95% CI)</td>
<td>6.7 (0.3–13.1)</td>
<td>1.4 (−4.8 to 7.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.028</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDAI ≤3.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 12</td>
<td>12.8 (9.7–16.0)</td>
<td>9.4 (6.7–12.1)</td>
<td>6.8 (3.9–9.7)</td>
<td>2.9 (1.3–4.6)</td>
</tr>
<tr>
<td></td>
<td>Treatment difference vs PBO (95% CI)</td>
<td>Exploratory p vs PBO</td>
<td>Exploratory p vs ADA</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>23.2 (19.3–27.1)</td>
<td>17.5 (14.0–21.0)</td>
<td>16.9 (12.7–21.2)</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs PBO</td>
<td>16.0 (11.3–20.7)</td>
<td>10.3 (6.0–14.7)</td>
<td>24.0 (19.2–28.8)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>29.7 (25.5–33.9)</td>
<td>24.6 (20.6–28.5)</td>
<td>16.9 (12.7–21.2)</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA</td>
<td>5.7 (−0.8 to 12.1)</td>
<td>0.6 (−5.7 to 6.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.064</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Boolean remission**

<table>
<thead>
<tr>
<th></th>
<th>Treatment difference vs PBO (95% CI)</th>
<th>Exploratory p vs PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 12</td>
<td>9.5 (6.7–12.2)</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs PBO</td>
<td>7.6 (4.5–10.7)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 24</td>
<td>18.5 (14.9–22.1)</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs PBO</td>
<td>13.7 (9.5–17.9)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>% (95% CI), week 52</td>
<td>22.5 (18.7–26.4)</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA</td>
<td>5.6 (−0.2 to 11.4)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.044</td>
<td></td>
</tr>
</tbody>
</table>

Comparisons vs ADA were for superiority unless otherwise indicated.

*Post hoc exploratory analysis.

ACR20, 20% improvement from baseline in American College of Rheumatology core criteria; ADA, adalimumab; CDAI, Clinical Disease Activity Index; CI, confidence interval; DAS28(CRP), Disease Activity Score in 28 joints with C-reactive protein; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; HAQ-DI, Health Assessment Questionnaire-Disability Index; LSM, least squares mean; N/A, not applicable; PBO, placebo; SD, standard deviation; SDAI, Simplified Disease Activity Index; SF-36 PCS, Short Form-36 Physical Component Summary.
## Table S4. Radiographic progression from baseline through week 52

<table>
<thead>
<tr>
<th>Campaign A</th>
<th>FIL200 (n = 475)</th>
<th>FIL100 (n = 480)</th>
<th>ADA (n = 325)</th>
<th>PBO (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients evaluated, n</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>467</td>
<td>471</td>
<td>319</td>
<td>466</td>
</tr>
<tr>
<td>Change from baseline at week 12</td>
<td>426</td>
<td>438</td>
<td>282</td>
<td>408</td>
</tr>
<tr>
<td>Change from baseline at week 24</td>
<td>405</td>
<td>404</td>
<td>271</td>
<td>351</td>
</tr>
<tr>
<td><strong>mTSS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline, mean ± SD</td>
<td>32.5 ± 48.0</td>
<td>36.7 ± 53.1</td>
<td>34.8 ± 55.0</td>
<td>31.6 ± 53.2</td>
</tr>
<tr>
<td>Change from baseline at week 12</td>
<td>0.08 ± 0.08</td>
<td>0.12 ± 0.08</td>
<td>0.13 ± 0.09</td>
<td>0.25 ± 0.09</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.18 (−0.30 to −0.06)</td>
<td>−0.14 (−0.25 to −0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>0.004</td>
<td>0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change from baseline at week 24</td>
<td>0.13 ± 0.09</td>
<td>0.15 ± 0.09</td>
<td>0.19 ± 0.10</td>
<td>0.40 ± 0.09</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.27 (−0.43 to −0.12)</td>
<td>−0.25 (−0.40 to −0.10)</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>p vs PBO</td>
<td>&lt;0.001</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0.05 (−0.22 to 0.12)</td>
<td>−0.03 (−0.20 to 0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.54</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Erosion score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline, mean ± SD</td>
<td>13.9 ± 24.2</td>
<td>16.8 ± 27.3</td>
<td>15.2 ± 28.7</td>
<td>14.0 ± 28.1</td>
</tr>
<tr>
<td>Change from baseline at week 12</td>
<td>0.03 ± 0.05</td>
<td>0.05 ± 0.05</td>
<td>0.05 ± 0.05</td>
<td>0.14 ± 0.05</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.11 (−0.18 to −0.05)</td>
<td>−0.10 (−0.16 to −0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change from baseline at week 24</td>
<td>0.03 ± 0.05</td>
<td>0.06 ± 0.05</td>
<td>0.06 ± 0.06</td>
<td>0.22 ± 0.05</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.19 (−0.27 to −0.10)</td>
<td>−0.16 (−0.24 to −0.07)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Joint space narrowing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline, mean ± SD</td>
<td>18.5 ± 25.6</td>
<td>19.9 ± 27.3</td>
<td>19.6 ± 28.2</td>
<td>17.6 ± 26.9</td>
</tr>
<tr>
<td>Change from baseline at week 12</td>
<td>0.05 ± 0.06</td>
<td>0.07 ± 0.06</td>
<td>0.09 ± 0.06</td>
<td>0.11 ± 0.06</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.06 (−0.15 to 0.02)</td>
<td>−0.04 (−0.12 to 0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>0.14</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change from baseline at week 24</td>
<td>0.10 ± 0.06</td>
<td>0.09 ± 0.06</td>
<td>0.12 ± 0.07</td>
<td>0.19 ± 0.07</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.09 (−0.19 to 0.02)</td>
<td>−0.09 (−0.20 to 0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>0.11</td>
<td>0.078</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Combined Campaign A and Campaign B

Patients evaluated at week 52, n

<table>
<thead>
<tr>
<th>mTSS</th>
<th>417</th>
<th>411</th>
<th>273</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change from baseline at week 24, LSM ± SE</td>
<td>−0.10 ± 0.10</td>
<td>0.19 ± 0.10</td>
<td>0.17 ± 0.11</td>
<td>0.49 ± 0.10</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−0.39 (−0.54 to −0.24)</td>
<td>−0.30 (−0.45 to −0.15)</td>
<td>Exploratory p vs PBO &lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0.07 (−0.23 to 0.09)</td>
<td>0.02 (−0.14 to 0.18)</td>
<td>Exploratory p vs ADA 0.40</td>
<td>0.80</td>
</tr>
<tr>
<td>Change from baseline at week 52, LSM ± SE</td>
<td>0.18 ± 0.12</td>
<td>0.45 ± 0.12</td>
<td>0.61 ± 0.13</td>
<td>N/A</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0.43 (−0.66 to −0.20)</td>
<td>−0.15 (−0.38 to 0.08)</td>
<td>Exploratory p vs ADA &lt;0.001</td>
<td>0.19</td>
</tr>
</tbody>
</table>

### Erosion score

| Change from baseline at week 24, LSM ± SE | 0.04 ± 0.06 | 0.08 ± 0.06 | 0.06 ± 0.06 | 0.27 ± 0.06 |
| Treatment difference vs PBO, LSM (95% CI) |−0.22 (−0.31 to −0.14) | −0.18 (−0.27 to −0.10) | Exploratory p vs PBO <0.001 | <0.001 |
| Treatment difference vs ADA, LSM (95% CI) |−0.01 (−0.11 to 0.08) | 0.03 (−0.07 to 0.12) | Exploratory p vs ADA 0.76 | 0.57 |
| Change from baseline at week 52, LSM ± SE | 0.06 ± 0.07 | 0.19 ± 0.07 | 0.28 ± 0.08 | N/A     |
| Treatment difference vs ADA, LSM (95% CI) |−0.22 (−0.35 to −0.09) | −0.09 (−0.23 to 0.04) | Exploratory p vs ADA 0.001 | 0.18 |

### Joint space narrowing

| Change from baseline at week 24, LSM ± SE | 0.06 ± 0.06 | 0.11 ± 0.06 | 0.11 ± 0.07 | 0.23 ± 0.07 |
| Treatment difference vs PBO, LSM (95% CI) |−0.17 (−0.26 to −0.08) | −0.12 (−0.22 to −0.03) | Exploratory p vs PBO <0.001 | 0.012 |
| Treatment difference vs ADA, LSM (95% CI) |−0.05 (−0.16 to 0.05) | −0.01 (−0.11 to 0.10) | Exploratory p vs ADA 0.29 | 0.90 |
| Change from baseline at week 52, LSM ± SE | 0.12 ± 0.07 | 0.26 ± 0.07 | 0.32 ± 0.08 | N/A     |
| Treatment difference vs ADA, LSM (95% CI) |−0.21 (−0.34 to −0.07) | −0.06 (−0.19 to 0.07) | Exploratory p vs ADA 0.002 | 0.36 |

*Patient numbers are not reported for the combined analyses due to the change from baseline including both Campaign A and Campaign B. ADA, adalimumab; CI, confidence interval; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; LSM, least squares mean; mTSS, van der Heijde modified total Sharp score; N/A, not applicable; PBO, placebo; SD, standard deviation; SE, standard error.
Table S5. Change from baseline in other ACR and DAS28(CRP) core measure components at weeks 12, 24, and 52

<table>
<thead>
<tr>
<th></th>
<th>FIL200</th>
<th>FIL100</th>
<th>ADA</th>
<th>PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 475)</td>
<td>(n = 480)</td>
<td>(n = 325)</td>
<td>(n = 475)</td>
</tr>
<tr>
<td><strong>TJC68</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>458</td>
<td>458</td>
<td>311</td>
<td>435</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−17 ± 11.1</td>
<td>−15 ± 10.7</td>
<td>−15 ± 9.9</td>
<td>−13 ± 11.6</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−4 (−5 to −2)</td>
<td>−3 (−4 to −1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−3 to −0)</td>
<td>−0 (−2 to 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.015</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>418</td>
<td>423</td>
<td>283</td>
<td>375</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−20 ± 12.1</td>
<td>−19 ± 10.9</td>
<td>−18 ± 11.1</td>
<td>−17 ± 11.7</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−3 (−4 to −2)</td>
<td>−2 (−3 to −1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−2 (−3 to −1)</td>
<td>−1 (−2 to 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.001</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>400</td>
<td>397</td>
<td>265</td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−21 ± 12.2</td>
<td>−21 ± 11.4</td>
<td>−20 ± 11.4</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−2 to −0)</td>
<td>−0 (−1 to 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.015</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TJC28</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>458</td>
<td>458</td>
<td>311</td>
<td>435</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−10 ± 6.4</td>
<td>−9 ± 6.2</td>
<td>−9 ± 5.7</td>
<td>−8 ± 6.7</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−2 (−3 to −2)</td>
<td>−1 (−2 to −1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−2 to −0)</td>
<td>−0 (−1 to 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.026</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>418</td>
<td>423</td>
<td>283</td>
<td>375</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>−12 ± 6.2</td>
<td>−12 ± 6.1</td>
<td>−11 ± 6.1</td>
<td>−11 ± 6.2</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−2 (−3 to −2)</td>
<td>−2 (−2 to −1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exploratory p vs PBO</td>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>Exploratory p vs ADA</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------</td>
<td>------------------------------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>400</td>
<td>−1 (−2 to −0)</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>13 ± 6.0</td>
<td>−12 ± 6.0</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0 (−1 to 0)</td>
<td>−12 ± 5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.012</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Exploratory p vs ADA</th>
<th>Treatment difference vs ADA, LSM (95% CI)</th>
<th>Exploratory p vs ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SJC66</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 12</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>458</td>
<td>−11 ± 7.5</td>
<td>0.060</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>11 ± 7.5</td>
<td>−12 ± 8.1</td>
<td>0.38</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−2 (−3 to −1)</td>
<td>−11 ± 7.1</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−2 to 0)</td>
<td>−0 (−1 to 0)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.004</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td><strong>Week 24</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>418</td>
<td>−13 ± 7.8</td>
<td>0.018</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>13 ± 7.4</td>
<td>−13 ± 7.4</td>
<td>0.60</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−2 (−3 to −1)</td>
<td>−12 ± 7.7</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−2 to 0)</td>
<td>−0 (−1 to 0)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.004</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>400</td>
<td>−14 ± 8.1</td>
<td>0.018</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>14 ± 8.1</td>
<td>−13 ± 7.6</td>
<td>0.60</td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−1 (−1 to 0)</td>
<td>−14 ± 7.5</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.018</td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Exploratory p vs ADA</th>
<th>Treatment difference vs ADA, LSM (95% CI)</th>
<th>Exploratory p vs ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SJC28</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 12</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>458</td>
<td>−8 ± 5.2</td>
<td>0.088</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>8 ± 5.4</td>
<td>−8 ± 5.4</td>
<td>0.60</td>
</tr>
<tr>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>−2 (−2 to −1)</td>
<td>−8 ± 4.9</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs PBO</td>
<td>0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>−0 (−1 to 0)</td>
<td>−0 (−1 to 0)</td>
<td></td>
</tr>
<tr>
<td>Exploratory p vs ADA</td>
<td>0.088</td>
<td>0.60</td>
<td></td>
</tr>
</tbody>
</table>
### SGA, mm

<table>
<thead>
<tr>
<th>Week</th>
<th>N</th>
<th>Mean ± SD</th>
<th>Treatment difference vs PBO, LSM (95% CI)</th>
<th>Exploratory p vs PBO</th>
<th>Treatment difference vs ADA, LSM (95% CI)</th>
<th>Exploratory p vs ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>400</td>
<td>−10 ± 5.1</td>
<td>−10 (−1 to −0)</td>
<td>&lt;0.010</td>
<td>−10 (−0 to 0)</td>
<td>0.010</td>
</tr>
</tbody>
</table>

### PGA, mm

<table>
<thead>
<tr>
<th>Week</th>
<th>N</th>
<th>Mean ± SD</th>
<th>Treatment difference vs PBO, LSM (95% CI)</th>
<th>Exploratory p vs PBO</th>
<th>Treatment difference vs ADA, LSM (95% CI)</th>
<th>Exploratory p vs ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>400</td>
<td>−44 ± 24.4</td>
<td>−42 (−12 to −4)</td>
<td>0.28</td>
<td>−42 (−42 to −4)</td>
<td>0.75</td>
</tr>
<tr>
<td>Week</td>
<td>N</td>
<td>Mean ± SD</td>
<td>Treatment difference vs PBO, LSM (95% CI)</td>
<td>Exploratory p vs PBO</td>
<td>Treatment difference vs ADA, LSM (95% CI)</td>
<td>Exploratory p vs ADA</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------------</td>
<td>------------------------------------------</td>
<td>----------------------</td>
<td>------------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td><strong>Week 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>413</td>
<td>−48 ± 19.2</td>
<td>−8 (−10 to −6)</td>
<td>&lt;0.001</td>
<td>−7 (−10 to −5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>419</td>
<td>−46 ± 19.6</td>
<td>−3 (−5 to −0)</td>
<td>0.030</td>
<td>−2 (−5 to 1)</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>283</td>
<td>−47 ± 19.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>373</td>
<td>−42 ± 20.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>−53 ± 18.2</td>
<td>−2 (−4 to 0)</td>
<td>0.11</td>
<td>1 (−3 to 1)</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>398</td>
<td>−50 ± 19.2</td>
<td>−1 (−3 to 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>265</td>
<td>−52 ± 18.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patient pain assessment, mm</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>457</td>
<td>−31 ± 26.9</td>
<td>−12 (−15 to −9)</td>
<td>&lt;0.001</td>
<td>−10 (−13 to −7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>458</td>
<td>−29 ± 25.3</td>
<td>−4 (−7 to −1)</td>
<td>0.018</td>
<td>−2 (−5 to 2)</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>311</td>
<td>−27 ± 23.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>435</td>
<td>−21 ± 26.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>418</td>
<td>−38 ± 27.0</td>
<td>−11 (−14 to −7)</td>
<td>&lt;0.001</td>
<td>−9 (−12 to −6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>423</td>
<td>−37 ± 25.6</td>
<td>−3 (−6 to 1)</td>
<td>0.096</td>
<td>−2 (−5 to 2)</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>283</td>
<td>−35 ± 24.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>376</td>
<td>−30 ± 27.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week 52</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>−43 ± 26.2</td>
<td>−2 (−5 to 2)</td>
<td>0.31</td>
<td>−0 (−4 to 3)</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>398</td>
<td>−41 ± 25.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>265</td>
<td>−41 ± 25.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### hsCRP, mg/L

<table>
<thead>
<tr>
<th>Week</th>
<th>N</th>
<th>Mean ± SD</th>
<th>Treatment difference vs PBO, LSM (95% CI)</th>
<th>Exploratory p vs PBO</th>
<th>Treatment difference vs ADA, LSM (95% CI)</th>
<th>Exploratory p vs ADA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 12</td>
<td>456</td>
<td>−11.0 ± 18.7</td>
<td>−8.0 (−9.9 to −6.1)</td>
<td>&lt;0.001</td>
<td>−1.8 (−3.8 to 0.3)</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>454</td>
<td>−9.6 ± 21.3</td>
<td>−6.5 (−8.3 to −4.6)</td>
<td>&lt;0.001</td>
<td>−0.2 (−2.3 to 1.9)</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>308</td>
<td>−7.8 ± 20.6</td>
<td>−3.3 (−2.7 to 0.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>431</td>
<td>−3.3 ± 22.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>416</td>
<td>−11.8 ± 20.7</td>
<td>−7.9 (−9.9 to −5.9)</td>
<td>&lt;0.001</td>
<td>−4.5 (−6.6 to −2.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>419</td>
<td>−10.5 ± 22.2</td>
<td>−6.6 (−8.6 to −4.6)</td>
<td>&lt;0.001</td>
<td>−3.2 (−5.3 to −1.0)</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>281</td>
<td>−6.2 ± 24.2</td>
<td>−4.0 (−5.9 to −2.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>370</td>
<td>−4.0 ± 19.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 52</td>
<td>396</td>
<td>−12.2 ± 20.8</td>
<td>−12.2 ± 20.8</td>
<td></td>
<td>−0.4 (−1.8 to 0.9)</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>386</td>
<td>−11.3 ± 23.1</td>
<td>−11.3 ± 23.1</td>
<td></td>
<td>0.8 (−0.5 to 2.2)</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>259</td>
<td>−9.6 ± 16.5</td>
<td>−9.6 ± 16.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Missing change scores were not imputed.

Comparisons vs ADA were for superiority.

ACR, American College of Rheumatology; ADA, adalimumab; CI, confidence interval; DAS28(CRP), Disease Activity Score in 28 joints with C-reactive protein; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; hsCRP, high-sensitivity C-reactive protein; LSM, least squares mean; PBO, placebo; PGA, Physician Global Assessment; SD, standard deviation; SGA, Subject Global Assessment; SJC28, swollen joint count in 28 joints; SJC66, swollen joint count in 66 joints; TJC28, tender joint count in 28 joints; TJC68, tender joint count in 68 joints.
Table S6. Proportions of patients achieving ACR20 at week 12 by region grouped by standard of care

<table>
<thead>
<tr>
<th>Group</th>
<th>FIL200</th>
<th>FIL100</th>
<th>ADA</th>
<th>PBO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>77/108 (71.3)</td>
<td>66/104 (63.5)</td>
<td>46/73 (63.0)</td>
<td>39/105 (37.1)</td>
</tr>
<tr>
<td>Group B</td>
<td>199/259 (76.8)</td>
<td>193/267 (72.3)</td>
<td>132/175 (75.4)</td>
<td>146/261 (55.9)</td>
</tr>
<tr>
<td>Group C</td>
<td>40/48 (83.3)</td>
<td>37/50 (74.0)</td>
<td>28/35 (80.0)</td>
<td>29/49 (59.2)</td>
</tr>
<tr>
<td>Group D</td>
<td>17/20 (85.0)</td>
<td>12/18 (66.7)</td>
<td>8/14 (57.1)</td>
<td>9/22 (40.9)</td>
</tr>
<tr>
<td>Group E</td>
<td>31/40 (77.5)</td>
<td>27/41 (65.9)</td>
<td>15/28 (53.6)</td>
<td>14/38 (36.8)</td>
</tr>
</tbody>
</table>

Data shown as n/N (%).

Group A, United States of America, Australia, New Zealand, Republic of Korea, Belgium, Germany, Italy, United Kingdom, South Africa, Canada, Spain, Ireland, Israel, and The Netherlands; Group B, Bulgaria, Czech Republic, Hungary, Slovakia, India, Ukraine, Poland, Romania, Russia, and Serbia; Group C, Argentina and Mexico; Group D, Hong Kong, Thailand, and Taiwan; Group E, Japan.

See Figure S2 for enrolment by country within each group.

ACR20, 20% improvement from baseline by American College of Rheumatology core criteria; ADA, adalimumab; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; PBO, placebo.
Table S7. ACR20, ACR50, and ACR70 response rates from Figure 2

<table>
<thead>
<tr>
<th>Week</th>
<th>FIL200 (n = 475)</th>
<th>FIL100 (n = 480)</th>
<th>ADA (n = 325)</th>
<th>PBO (n = 475)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACR20</td>
<td>ACR50</td>
<td>ACR70</td>
<td>ACR20</td>
</tr>
<tr>
<td>2</td>
<td>37.3</td>
<td>9.1</td>
<td>2.7</td>
<td>27.5</td>
</tr>
<tr>
<td>4</td>
<td>51.6</td>
<td>22.3</td>
<td>9.1</td>
<td>45.6</td>
</tr>
<tr>
<td>8</td>
<td>69.7</td>
<td>35.6</td>
<td>17.3</td>
<td>63.3</td>
</tr>
<tr>
<td>12</td>
<td>76.6</td>
<td>47.2</td>
<td>26.1</td>
<td>69.8</td>
</tr>
<tr>
<td>14</td>
<td>81.5</td>
<td>52.8</td>
<td>32.4</td>
<td>76.9</td>
</tr>
<tr>
<td>16</td>
<td>80.2</td>
<td>55.4</td>
<td>33.9</td>
<td>76.7</td>
</tr>
<tr>
<td>20</td>
<td>81.5</td>
<td>55.6</td>
<td>34.1</td>
<td>77.7</td>
</tr>
<tr>
<td>24</td>
<td>78.1</td>
<td>57.9</td>
<td>36.2</td>
<td>77.7</td>
</tr>
<tr>
<td>26</td>
<td>77.9</td>
<td>59.8</td>
<td>38.1</td>
<td>75.8</td>
</tr>
<tr>
<td>30</td>
<td>80.2</td>
<td>60.4</td>
<td>38.7</td>
<td>77.5</td>
</tr>
<tr>
<td>36</td>
<td>79.6</td>
<td>61.3</td>
<td>39.8</td>
<td>76.5</td>
</tr>
<tr>
<td>44</td>
<td>79.4</td>
<td>62.1</td>
<td>42.1</td>
<td>77.7</td>
</tr>
<tr>
<td>52</td>
<td>78.3</td>
<td>62.3</td>
<td>44.2</td>
<td>75.6</td>
</tr>
</tbody>
</table>

ACR20/50/70, 20%/50%/70% improvement from baseline based on American College of Rheumatology Criteria; ADA, adalimumab; FIL100, filgotinib 100 mg; FIL200, filgotinib 200 mg; PBO, placebo.
**Figure S1. Hierarchical testing of secondary endpoints**

| 1) Superiority of filgotinib 100 mg compared with placebo based on ACR20 response rate at week 12 | P-value in trial |
| 2) Superiority of filgotinib 200 mg compared with placebo based on the change from baseline in HAQ-DI at week 12 | 0.001 |
| 3) Superiority of filgotinib 100 mg compared with placebo based on the change from baseline in HAQ-DI at week 12 | 0.001 |
| 4) Superiority of filgotinib 200 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.001 |
| 5) Superiority of filgotinib 200 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.001 |
| 8) Non-inferiority of filgotinib 200 mg compared with adalimumab based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.001 |
| 9) Non-inferiority of filgotinib 100 mg compared with adalimumab based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.054 |
| 10) Superiority of filgotinib 100 mg compared with placebo based on the change from baseline in SF-36 PCS at week 12 | 0.001 |
| 11) Superiority of filgotinib 100 mg compared with placebo based on the change from baseline in SF-36 PCS at week 12 | 0.001 |
| 12) Superiority of filgotinib 200 mg compared with placebo based on the change from baseline in FACIT-Fatigue at week 12 | 0.001 |
| 13) Superiority of filgotinib 100 mg compared with placebo based on the change from baseline in FACIT Fatigue at week 12 | 0.001 |
| 14) Superiority of filgotinib 200 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.009 |
| 15) Superiority of filgotinib 100 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.009 |
| 16) Non-inferiority of filgotinib 200 mg compared with adalimumab based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.005 |
| 17) Non-inferiority of filgotinib 100 mg compared with adalimumab based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.002 |
| 18) Superiority of filgotinib 200 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.001 |
| 19) Superiority of filgotinib 100 mg compared with placebo based on the proportion of subjects with DAS28(CRP)<2.6 at week 12 | 0.001 |

First non-significant result shown in red. Subsequent results not adjusted for multiplicity are shaded gray.

*Exploratory p value.

ACR20, 20% improvement from baseline in American College of Rheumatology core criteria;

DAS28(CRP), Disease Activity Score in 28 joints with C-reactive protein; FACT-F, Functional Assessment of Chronic Illness Therapy; HAQ-DI, Health Assessment Questionnaire-Disability Index;
mTSS, van der Heijde modified total Sharp score; SF-36, Short Form-36 Physical Component Summary.
Figure S2. Treated patients by country of enrollment and group

![Bar chart showing treated patients by country and group]

Group A = 390 (22.2%)

Group B + C + D + E = 1365 (77.8%)
**Figure S3.** Radiographic progression through week 52. **A)** mTSS change from baseline, **B)** erosion score change from baseline, and **C)** joint space narrowing change from baseline
Combined data from Campaign A (through week 24) and Campaign B (through week 52 including rereading of baseline and week 24) are shown. Supporting values are shown in Table S4.

Error bars represent the SE of the LS mean.

*Exploratory $p < 0.05$, ***exploratory $p < 0.001$ vs PBO.

++Exploratory $p < 0.01$, +++exploratory $p < 0.001$ vs ADA.

ADA, adalimumab; FIL, filgotinib; LS, least-squares; mTSS, van der Heijde modified total Sharp score;
PBO, placebo; SE, standard error.