Protocol !4V-MC-JAHO (e)

A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata.

NCT03570749

Approval Date: 02-Nov-2020

Protocol I4V-MC-JAHO(e)
A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata

BRAVE-AA1

Baricitinib (LY3009104)

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol Electronically Signed and Approved by Lilly on 21-May-2018
Amendment (a) Electronically Signed and Approved by Lilly on 28-Jun-2018
Amendment (b) Electronically Signed and Approved by Lilly on 30-Oct-2018
Amendment (c) Electronically Signed and Approved by Lilly on 18-Oct-2019
Amendment (d) Electronically Signed and Approved by Lilly on 11-May-2020
Amendment (e) Electronically Signed and Approved by Lilly on approval date provided below.

Approval Date: 02-Nov-2020 GMT

Table of Contents

Section **Page**

List of Tables

Table **Page**

List of Figures

Figure **Page**

List of Appendices

Appendix **Page**

1. Synopsis

Title of Study:
A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata.

Rationale:
Although, in the past, alopecia areata (AA) has been considered by some to be a benign condition, severe AA is now recognized as a significant medical condition with emotional and psychosocial distress, including high prevalence of depression and anxiety (Colon et al. 1991; Hunt and McHale 2005; Villasante Fricke and Miteva 2015). Additionally, reduction in health-related quality of life experienced by patients with AA has been well-documented (Jankovic et al. 2016; Liu et al. 2016; Rencz et al. 2016). Alopecia areata can be reversible – either spontaneously or after various types of treatments, but results are inconsistent (Islam et al. 2015) and there is currently no Food and Drug Administration- (FDA-) approved treatment for AA.

Baricitinib is an orally available, selective inhibitor of Janus kinases (JAKs). Janus kinases are a family of tyrosine kinases that mediate cytokine receptor signaling through phosphorylation and activation of signal transducers and activators of transcription (STAT) proteins. There are 4 known JAK family members: JAK1, JAK2, JAK3, and tyrosine kinase 2 (TYK2) (Clark et al. 2014). The relative affinity of JAK inhibitors for different members of the JAK kinase family allows for differentiation of JAK inhibitors in relation to their enzymatic inhibitory profile. In vitro assays indicate that baricitinib is a selective inhibitor of JAKs with potency and selectivity for JAK1 and JAK2 and less potency for JAK3 or TYK2 (Fridman et al. 2010).

Dual inhibition of JAK1 and JAK2, which may interrupt interferon gamma (IFNγ) signaling and other inflammatory pathways that contribute to the immunopathogenesis of AA, and clinical evidence with other JAK inhibitors support the investigation of baricitinib in the treatment of AA. This adaptive, operationally seamless, Phase 2/3, placebo-controlled study is designed to select up to 2 doses of baricitinib and assess their efficacy and safety for the treatment of severe (Severity of Alopecia Tool [SALT] score of 50% to 94%) and very severe AA (SALT score of 95% to 100%). Additionally, it will include a randomized withdrawal to explore the persistence of treatment effects and efficacy of retreatment upon relapse. This operationally seamless design will provide longer-term safety data earlier in the clinical development program, compared to running separate Phase 2 and Phase 3 studies.

Results from the interim analysis of the Phase 2 portion of Study JAHO showed that the 2-mg and 4-mg baricitinib doses demonstrated numerical superiority over placebo and the 1-mg dose in hair regrowth after 12 ($SALT_{30}$) and 16 ($SALT_{50}$) weeks of treatment with no new safety signal. These results have been confirmed by a second interim analysis of Phase 2 portion of Study JAHO where the 2-mg dose and 4-mg dose were statistically significantly superior to placebo on the primary endpoint ($SALT \leq 20$) after 36 weeks of treatment.

Objective(s)/Endpoints:

Objectives	Endpoints
Primary Objective	
To test the hypothesis that the 4-mg dose or 2-mg dose of baricitinib is superior to placebo in the treatment of patients with severe or very severe AA	• Proportion of patients achieving SALT ≤20 at Week 36.
Key Secondary Objectives (Double–Blind, Placebo-Controlled Treatment Period) *These are pre-specified objectives that will be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as measured by **physician-assessed** signs and symptoms of AA	• Proportion of patients achieving SALT ≤20 at Weeks 16 and 24 • Percent change from baseline in SALT score at Week 36. • Proportion of patients achieving a $SALT_{50}$ at Week 12. • Proportion of patients achieving a $SALT_{90}$ at Week 36. • Proportion of patients achieving an absolute SALT ≤10 at Weeks 24 and 36. • Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week 36 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline) • Proportion of patients achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week 36 (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline).
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as assessed by a **PRO measure**	• Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement from Baseline at Week 36 among patients with a score of ≥3 at Baseline.
Other Secondary Objectives (Double–Blind, Placebo-Controlled Treatment Period) *These are pre-specified objectives that will NOT be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as measured by **physician-assessed** signs and symptoms of AA	• Proportion of patients achieving $SALT_{50}$ at Weeks 16, 24, and 36. • Proportion of patients achieving a $SALT_{75}$ at Weeks 24 and 36. • Proportion of patients achieving a $SALT_{90}$ at Week 24 • Change from Baseline in SALT score at Weeks 12, 16, 24, and 36. • Percent change from Baseline in SALT score at Weeks 12, 16, and 24. • Time to achieve SALT ≤20. • Proportion of patients achieving $SALT_{100}$ at Weeks 24 and 36. • Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Weeks 16 and 24 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline). • Proportion of patients achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from

Objectives	Endpoints
	baseline at Weeks 16 and 24 (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline).
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as assessed by **PRO measures** and quality of life tools	• Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement from Baseline at Weeks 12 and 24 among patients with a score of ≥3 at Baseline • Proportion of patients achieving PRO Measure for EB 0 or 1 with ≥2-point improvement from baseline at Weeks 16, 24, and 36 (among patients with PRO Measure for EB ≥2 at Baseline) • Proportion of patients achieving PRO Measure for EL 0 or 1 with ≥2-point improvement from baseline at Weeks 16, 24, and 36 (among patients with PRO Measure for EL ≥2 at Baseline) • Mean change from Baseline in HADS-A and HADS-D total scores at Weeks 24 and 36
Other Secondary (Patients entering Randomized Withdrawal) *These are pre-specified objectives that will NOT be adjusted for multiplicity.*	
To compare the maintenance of efficacy for patients randomized to remain on baricitinib, compared with patients randomized to placebo at Week 52 of the long-term extension period, as measured by **physician-assessed** signs of AA	• Proportion of patients maintaining SALT ≤20 at Weeks 64, 76, 88, 104, 120, 136, 152, 168, 184, and 200. • Proportion of patients experiencing a loss of treatment benefit (>20-point absolute worsening in SALT score) at Weeks 64, 76, 88, 104, 120, 136, 152, 168, 184, and 200. • Time to loss of treatment benefit (>20-point absolute worsening in SALT score).
For patients experiencing loss of treatment benefit after randomization to placebo at Week 52: • To evaluate the recapture of efficacy for patients who were retreated after experiencing a loss of treatment benefit during the long-term maintenance period as measured by **physician-assessed** signs of AA • To evaluate the recapture of efficacy for patients who were retreated after experiencing a loss of treatment benefit during the long-term maintenance period as assessed by **PRO** and quality of life tools	• Proportion of patients that achieve a SALT ≤20 at 12, 16, 24, and 36 weeks of retreatment with baricitinib. • Percent change in SALT score at 12, 16, 24, and 36 weeks of retreatment with baricitinib. • Proportion of patients with a PRO for Scalp Hair Assessment score of 0 or 1 at 12, 16, 24, and 36 weeks of retreatment with baricitinib.
Exploratory Objectives may include evaluating the response to baricitinib treatment regimens on clinical measures and patient-reported outcomes. These endpoints may include dichotomous endpoints or change from baseline for the following measures: SALT, AA-IGA, $SALT_{30,}$ ClinROs for Nail Appearance, Eyebrows, and/or Eyelash Hair Loss, PROs for Scalp Hair Assessment, Eyebrows, and Eyelashes, Nail Appearance and Eye Irritation, Skindex-16 AA, SF-36, EQ-5D-5L, HADS. Assessments of efficacy may be performed beyond Week 104 up to Week 200. In addition, baricitinib pharmacokinetics will be characterized in the AA population and relationships between exposure and study endpoints will be explored.	

Abbreviations: AA = alopecia areata; AA-IGA = Alopecia Areata Investigator Global Assessment; ClinRO = clinician-reported outcome; EB = eyebrow; EL = eyelash; EQ-5D-5L = European Quality of Life – 5 Dimensions – 5 Level; HADS = Hospital Anxiety and Depression Scale; PRO = patient-reported outcome; SALT = Severity of Alopecia Tool; $SALT_{30}$ = at least 30% improvement from Baseline in SALT score; $SALT_{50}$ = at least 50% improvement from Baseline in SALT score; $SALT_{75}$ = at least 75% improvement from Baseline in SALT score; $SALT_{90}$ = at least 90% improvement from Baseline in SALT score; $SALT_{100}$ = 100% improvement from Baseline in SALT score; SF-36 = Short Form-36 Health Survey acute version 2; Skindex-16 AA = Skindex-16 Adapted for Alopecia Areata.

Summary of Study Design:

Study I4V-MC-JAHO (JAHO) is an adaptive, operationally seamless, Phase 2/3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study designed to identify up to 2 doses of baricitinib to be evaluated further in the Phase 3 portion of the study. In Phase 3, efficacy and safety of 2-mg once daily (QD) and 4-mg QD doses of baricitinib will be compared to placebo in adult patients with severe (SALT score of 50% to 94%) or very severe (SALT score of 95% to 100%) scalp AA. Approximately 725 adult patients will be enrolled into Study JAHO. Approximately 100 patients will be enrolled into the Phase 2 portion of the study and approximately 625 patients will be enrolled into the Phase 3 portion of the study.

Patients must have a current AA episode of more than 6 months' duration prior to screening (Visit 1), with at least 50% scalp involvement at screening AND Baseline (Visits 1 and 2) with no spontaneous improvement (no more than a 10 point reduction in SALT) over the past 6 months. Patients with a current episode of severe or very severe AA of more than 8 years will not be eligible for inclusion in the study unless episodes of regrowth, spontaneous or under treatment, have been observed on the affected areas of the scalp over the past 8 years.

Study Stages and Treatment Arms

The enrollment of patients in the study will be divided into 2 stages, which are separated by the Decision Point. Different randomization schemes at Baseline (Visit 2) will be used by the interactive web-response system (IWRS) (during Stage 1 [Phase 2], Stage 1 [Phase 3] and Stage 2).

- Stage 1: The time from study start until the Decision Point. A maximum of approximately 300 patients will be randomized during Stage 1, before the Decision Point. The first approximately 100 randomized patients will comprise the Phase 2 portion of the study and will be randomized in a 1:1:1:1 ratio to receive placebo QD, baricitinib 1-mg QD, baricitinib 2-mg QD, or baricitinib 4-mg QD. An interim analysis will be conducted when the first approximately 100 patients who have been randomized and received treatment have reached Week 12 or have discontinued prior to Week 12. The remaining approximately 200 patients enrolled during Stage 1 will contribute patients to the Phase 3 portion of the study and will be randomized at a 2:2:3 ratio to receive placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD.
- Decision Point: The point in time when up to 2 baricitinib doses will be selected to continue in Stage 2 or the study will be stopped for futility, based on the outcome of the interim analysis. After the Decision Point, up to 3 treatments arms will continue into Stage 2: placebo; baricitinib low dose (lowest dose remaining in the study); and baricitinib high dose (highest dose remaining in the study).

- Stage 2: The time after the Decision Point until the end of the study during which the remaining patients (approximately 425 patients) will be enrolled into the Phase 3 portion of the study and randomized at a ratio of 2:2:3 to receive placebo QD, baricitinib low dose QD, or baricitinib high dose QD.
- Based on the Week 12 interim analysis of the Phase 2 population of this study, baricitinib 2-mg QD (low dose) and baricitinib 4-mg QD (high dose) were selected at the Decision Point to continue into Stage 2 of Study JAHO.

Transitioning Patients After Decision Point

- After the Decision Point, patients who were enrolled in the baricitinib dose group that is discontinued (1-mg) will transition to the highest dose of baricitinib remaining in the study (4-mg).
- Patients and sites will remain blinded to treatment allocation after the Decision Point and, therefore, will not know which patients will be transitioned. Transition will automatically occur at the next visit after Decision Point; this will be referred to as the Transition Visit. A patient should be seen within 8 weeks following the Decision Point being communicated to the sites. If there is not a regularly scheduled visit during this timeframe, patients may be brought in for an unscheduled visit. A patient should be seen within 8 weeks following the Transition Visit to obtain laboratory values for safety review. If there is not a regularly scheduled visit during this timeframe, patients may be brought in for an unscheduled visit. After the Decision Point, all patients enrolled during Stage 1 will follow all protocol procedures for Periods 2, 3 and 4.

Number of Participants:

Planned enrollment is approximately 100 patients in the Phase 2 portion and approximately 625 patients in the Phase 3 portion. A maximum of approximately 300 patients will be enrolled during Stage 1 (approximately 100 patients in Phase 2 and approximately 200 patients in Phase 3). The remaining approximately 425 patients for Phase 3 will be enrolled during Stage 2.

Study Design

The study design includes 5 periods: a 5-week screening period; a 36-week double-blind, placebo-controlled treatment period; a 68-week long-term extension period; a 96-week bridging extension, and a posttreatment follow-up period. The Schedule of Activities (SOA) will be the same for patients randomized during Stage 1 and Stage 2, except that some patients randomized during Stage 1 might have one or more unscheduled visits after the Decision Point:

- **Period 1:** Screening period (Visit 1) is between 3 and 35 days prior to Visit 2 (Week 0).
- **Period 2:** 36-week double-blind, placebo-controlled treatment period is from Week 0 (Baseline; Visit 2) to Week 36 (Visit 8).
- **Period 3**: 68-week, long-term extension period with randomized withdrawal (for responders) is from Week 36 (Visit 8) to Week 104 (Visit 18).
- **Period 4**: Bridging Extension: from Week 104 (Visit 18) and up to Week 200 (Visit 24). Subjects who have completed Week 104 and have not met criteria of permanent

discontinuation will have the possibility to remain in the trial for up to 96 additional weeks (up to Week 200).

- **Period 5:** Posttreatment follow-up period; the posttreatment follow-up visit should occur approximately 4 weeks after the last dose of investigational product (IP) (i.e., 4-weeks after Week 200 (Visit 24) or the Early Termination Visit [ET]).

Note: Patients who have discontinued IP and remain in the study for more than 28 days without IP will have an ET; however, a separate follow-up visit (V801) is not required. Patients who have completed Week 200 and will continue on marketed product beyond Week 200 do not need to complete Period 5 (Visit 801).

At Week 36

The predicted enrollment rate for Study JAHO suggests that some of the patients enrolled during Stage 1 may reach Week 36 before the Decision Point is achieved:

- Patients in the placebo treatment arm who have not achieved a SALT ≤20 at Week 36 AND who reach the Week 36 Visit
  - prior to the Decision Point, will be rescued to baricitinib and randomized in a 1:1 ratio to baricitinib 2-mg or baricitinib 4-mg.
  - after the Decision Point, will be rescued to baricitinib and randomized in a 1:1 ratio to baricitinib 2-mg or baricitinib 4-mg.
- Patients in the placebo arm who have achieved a SALT ≤20 will remain on placebo at Week 36. These patients who have experienced spontaneous regrowth will remain on placebo for the remainder of the trial, even if relapse is observed later during the study.
- All patients in the baricitinib treatment arms will continue in their current treatment group regardless of their treatment response at Week 36 (refer to Week 52 and Week 76 sections below for additional rescue information and discontinuation criteria).

At Week 52

<u>Responders (</u>SALT ≤20)

- Patients in baricitinib treatment arms who achieve a SALT ≤20 at Week 52 (responders) are eligible for randomized withdrawal, provided that they have stayed on the same dose of baricitinib from initial randomization (Visit 2).
- Responders who have had a change in dose after the Decision Point, or patients who were rescued to baricitinib at Week 36, will not be eligible for randomized withdrawal and will remain in their same treatment group.
- Eligible patients will be automatically randomized in a blinded manner by the IWRS in a 3:1 ratio to either stay on their current dose of baricitinib or transition to placebo (randomized withdrawal).
- Any patients in the placebo treatment arm at Week 52 who have achieved a SALT ≤20 will remain on placebo.

- Responders who experience a loss of treatment response after Week 52 (defined as >20-point absolute worsening in total SALT score), and who:
    - were randomized to placebo at Week 52 (randomized withdrawal), will be automatically retreated with their baricitinib dose, as randomized at Baseline (Visit 2).
    - remained on baricitinib at Week 52 (randomized withdrawal), will continue to receive the same dose of baricitinib.
    - were randomized to placebo at Baseline (Visit 2), will remain on placebo.

Note: An unscheduled visit after Week 52 may be used to assess the patient who reports a loss of treatment response.

Nonresponders (SALT >20)

- Patients who have been in the baricitinib 4-mg treatment group from Baseline AND have never achieved a SALT ≤20 by Week 52 AND do not have a ≥2-point improvement from Baseline in clinician-reported outcome (ClinRO) measure for eyebrow or eyelash hair loss at Week 52 will be automatically transitioned to placebo.
- Patients who have been in the baricitinib 4-mg treatment group and have achieved a SALT ≤20 before Week 52 and have lost response, will remain on baricitinib 4-mg.
- Those who have been in the baricitinib 2-mg treatment group from Baseline will be rescued to baricitinib 4-mg.
- Those who were transitioned to baricitinib 4-mg after the Decision Point, or were rescued to baricitinib at Week 36, will continue in their current treatment arm at Week 52.

At Week 76

- Patients who are nonresponders, SALT >20, at Weeks 52 AND 76 will be automatically discontinued from the study at Week 76, unless they have a ≥2-point improvement from Baseline in ClinRO measure for eyebrow or eyelash hair loss.

Statistical Analysis:

Analysis to select up to 2 doses to advance to Stage 2 of the trial or stop the trial for futility will be conducted in the Phase 2, Week 12 Interim Analysis Set (IAS), which includes the first approximately 100 randomized and treated patients in Stage 1 who complete the Visit 5 (Week 12) assessment or discontinue early. A 2nd interim analysis will be conducted when all Phase 2 patients have completed 36 weeks of treatment or discontinued early and this efficacy analysis will use the Phase 2, Week 36 IAS. Unless otherwise specified, the efficacy and health outcome analyses during the Phase 3 portion of the study will be conducted on the Full Analysis Set (FAS), which includes all patients enrolled in Phase 3 portion and randomized to baricitinib 4-mg, baricitinib 2-mg, and placebo treatment arms in both Stages 1 and 2. Safety analyses will be conducted in safety population. Additional efficacy or safety interim analyses prior to the final database lock (F-DBL) may occur to support regulatory submissions and scientific disclosures and will be described in detail in the SAP.

Treatment comparisons of discrete efficacy variables will be made using a logistic regression analysis as primary analysis with geographic region, duration of current episode at Baseline (<4 years vs ≥4 years), baseline value, and treatment group in the model unless, otherwise, stated. In the case when logistic regression model does not produce statistical results due to sparse data, Fisher exact test will be used. The proportions and 95% confidence intervals (CIs) will be reported. Patients will be generally considered nonresponders for the nonresponder imputation- (NRI-) based analysis if they do not meet clinical response criteria or if they discontinued study or study treatment at any time prior to the timepoint of interest for any reason. Additional analyses may be performed and will be described in detail in the statistical analysis plan (SAP) for patients who discontinued the study and for patients missing measurements prior to a fixed timepoint.

Treatment comparisons of continuous efficacy and health outcome variables will be made using analysis of covariance (ANCOVA) with geographic region, duration of current episode at Baseline (<4 years vs ≥4 years), treatment group, and baseline value in the model unless otherwise stated. Type III tests for least square (LS) means will be used for the statistical comparison between treatment groups. The LS mean difference, standard error, p-value, and 95% CI will also be reported. Additional analyses may be performed and will be described in detail in the SAP.

Time-to-event analysis will be done and analyzed using log-rank test. A Cox proportional hazards model may be used with treatment and other stratification variables in the model unless otherwise stated. Hazard ratio with CIs may be reported. Kaplan–Meier curves will also be produced. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed and these would be described in detail in the SAP. If the assumption of proportional hazards is not justified, nonproportionality may be modeled by stratification.

Fisher exact test will be used for all adverse events (AEs), discontinuation, and other categorical safety data. Continuous vital signs and laboratory values will be analyzed by an ANCOVA with treatment and baseline values in the model.

2. Schedule of Activities

The Schedule of Activities described below should be followed for all participants enrolled in Study JAHO. In the event participation in this study is affected by exceptional circumstances (such as pandemics or natural disasters), please refer to Appendix 8 and consult with the sponsor's representative for additional guidance.

Table JAHO.1. Schedule of Activities

	Screen-ing	Double-blind Treatment Period							Long-term Extension										Bridging Extension							PTFU
	Period 1	Period 2 (36 weeks)							Period 3 (68 weeks)										Period 4 (96 weeks)							Period 5
Visit Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24/ ET	UV[a]	801
Weeks from Randomization		0	4	8	12	16	24	36	40	44	52	56	60	64	68	76	88	104	120	136	152	168	184	200	Period2 to 4	
Visit tolerance interval (days)	-3 to -35		±4	±4	±4	±4	±7	±7	±4	±4	±7	±4	±4	±4	±4	±7	±7	±7	±10	±10	±10	±10	±10	±10		28 ±4
Inclusion and Exclusion Review	X	X																								
Informed consent	X																									
Randomization		X						X[b]			X[b]															
Clinical Assessments																										
Demographics	X																									
Preexisting Conditions/Medical History	X																									
MPHL/FPHL History	X																									
Substance Use (alcohol, tobacco)	X																									
Previous and Current AA Treatments	X																									
Chest x-ray (posterior-anterior view)[c]	X																									
TB test[d]	X																									
Read PPD if applicable (48 to 72 hours after V1)[e]	X																									
Physical	X																									

	Screen-ing	Double-blind Treatment Period							Long-term Extension										Bridging Extension							PTFU
	Period 1	Period 2 (36 weeks)							Period 3 (68 weeks)										Period 4 (96 weeks)							Period 5
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24/ ET**	**UV[a]**	**801**
Weeks from Randomization		**0**	**4**	**8**	**12**	**16**	**24**	**36**	**40**	**44**	**52**	**56**	**60**	**64**	**68**	**76**	**88**	**104**	**120**	**136**	**152**	**168**	**184**	**200**	**Period2 to 4**	
Visit tolerance interval (days)	**-3 to -35**		**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±4**	**±4**	**±7**	**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±10**	**±10**	**±10**	**±10**	**±10**	**±10**		**28 ±4**
Examination																										
12-lead ECG (single, local)	X																									
Height	X																									
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
BMI	X	X																								
Vital signs (BP and Pulse)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
Symptom-directed Physical Exam[f]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
MPHL/FPHL Assessment[g]								X			X							X						X		
Adverse Events		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant Medication	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
IWRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Photography (4 planes of scalp and 1 frontal view of face and scalp)[h]		X			X			X			X													X[h]		
Photography Eyebrows/ Eyelashes[h]		X			X			X			X													X[h]		
Photography Nails (if involvement)[h]		X			X			X			X													X[h]		
IP dispensed		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	
IP returned and compliance			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

	Screen-ing	Double-blind Treatment Period							Long-term Extension										Bridging Extension							PTFU
	Period 1	Period 2 (36 weeks)							Period 3 (68 weeks)										Period 4 (96 weeks)							Period 5
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24/ ET**	**UV[a]**	**801**
Weeks from Randomization		**0**	**4**	**8**	**12**	**16**	**24**	**36**	**40**	**44**	**52**	**56**	**60**	**64**	**68**	**76**	**88**	**104**	**120**	**136**	**152**	**168**	**184**	**200**	**Period2 to 4**	
Visit tolerance interval (days)	**-3 to -35**		**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±4**	**±4**	**±7**	**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±10**	**±10**	**±10**	**±10**	**±10**	**±10**		**28 ±4**
assessed																										
Scales/ Questionnaires																										
SALT (AA-IGA)[i]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hamilton-Norwood Scale[g]	X							X			X							X						X		
ClinRO Measure for Eyebrow Hair Loss		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
ClinRO Measure for Eyelash Hair Loss		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
ClinRO Measure for Nail Appearance		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
PRO for Scalp Hair Assessment	X (Stage 2 only)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
PRO Measure for Eyebrows		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
PRO Measure for Eyelashes		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
PRO Measure for Eye Irritation		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
PRO Measure for Nail Appearance		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
EQ-5D-5L		X			X		X	X			X			X		X		X						X		X
SF-36		X			X		X	X			X			X		X		X						X		X
Skindex-16 Adapted for AA (Stage 2 only)		X			X		X	X			X			X		X		X						X		X

	Screen-ing	Double-blind Treatment Period							Long-term Extension										Bridging Extension							PTFU
	Period 1	Period 2 (36 weeks)							Period 3 (68 weeks)										Period 4 (96 weeks)							Period 5
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24/ ET**	**UV[a]**	**801**
Weeks from Randomization		**0**	**4**	**8**	**12**	**16**	**24**	**36**	**40**	**44**	**52**	**56**	**60**	**64**	**68**	**76**	**88**	**104**	**120**	**136**	**152**	**168**	**184**	**200**	**Period2 to 4**	
Visit tolerance interval (days)	**-3 to -35**		**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±4**	**±4**	**±7**	**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±10**	**±10**	**±10**	**±10**	**±10**	**±10**		**28 ±4**
HADS		X			X		X	X			X			X		X		X						X		X
C-SSRS and Self-Harm Supplement[j]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
Self-Harm Follow-up Form[k]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
Laboratory Assessment																										
Lipids (Fasting Visit)[l]		X			X		X	X			X			X		X		X	X	X	X	X	X	X		X
Clinical Chemistry[m]	X	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X	X	X	X	X	X	X		X
Hematology	X	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X	X	X	X	X	X	X		X
Serum Pregnancy Test[n]	X																									
FSH[o]	X																									
TSH	X																									
HIV	X																									
HCV antibody testing[p]	X																									
HBV testing[q]	X																									
HBV DNA[q]	X				X		X	X			X			X		X	X	X	X	X	X	X	X	X		X
Urinalysis	X	X	X	X	X	X	X	X			X	X	X	X	X		X	X						X		X
Urine Pregnancy Test[n]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
Pharmacogenetics: blood		X																								
Serum immunoglobulin		X			X			X			X															

	Screen-ing	Double-blind Treatment Period							Long-term Extension										Bridging Extension							PTFU
	Period 1	Period 2 (36 weeks)							Period 3 (68 weeks)										Period 4 (96 weeks)							Period 5
Visit Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24/ ET**	**UV[a]**	**801**
Weeks from Randomization		**0**	**4**	**8**	**12**	**16**	**24**	**36**	**40**	**44**	**52**	**56**	**60**	**64**	**68**	**76**	**88**	**104**	**120**	**136**	**152**	**168**	**184**	**200**	**Period2 to 4**	
Visit tolerance interval (days)	**-3 to -35**		**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±4**	**±4**	**±7**	**±4**	**±4**	**±4**	**±4**	**±7**	**±7**	**±7**	**±10**	**±10**	**±10**	**±10**	**±10**	**±10**		**28 ±4**
(IgE)																										
Exploratory storage samples (serum, plasma)		X			X		X	X			X			X		X		X						X		X
RNA and biomarkers: blood		X			X		X	X			X			X		X		X						X		X
Baricitinib plasma concentration (PK sample)[r]		X	X	X	X	X																		X[r]		

Abbreviations: AA = alopecia areata; AA-IGA = Alopecia Areata Investigator Global Assessment; AGA = androgenetic alopecia; BMI = body mass index; BP = blood pressure; CKD-EPI = Chronic Kidney Disease Epidemiology Collaboration; ClinRO = clinician-reported outcome; C-SSRS = Columbia Suicide Severity Rating Scale; DNA = deoxyribonucleic acid; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; EQ-5D-5L = European Quality of Life – 5 Dimensions – 5 Level; ET = early termination; FPHL = female pattern hair loss; FSH = follicle-stimulating hormone; HBcAb = hepatitis B core antibody; HBsAb = hepatitis B surface antibody; HADS = Hospital Anxiety and Depression Scale; HBV = hepatitis B virus; HCV = hepatitis C virus; HIV = human immunodeficiency virus; IgE = immunoglobulin E; IP = investigational product; IWRS = investigator web response system; MPHL = male pattern hair loss; PK = pharmacokinetic; PPD = purified protein derivative; PRO = patient-reported outcome; PTFU = posttreatment follow-up; RNA = ribonucleic acid; SALT = Severity of Alopecia Tool; SF-36 = Short Form-36 Health Survey acute version 2; TB = tuberculosis; TSH = thyroid-stimulating hormone; UV = unscheduled visit; V = Visit.

[a] Only applicable to patients enrolled during Stage 1, who will need an unscheduled visit(s) after the Decision Point to potentially transition to baricitinib 4-mg (SALT assessment not required) and to patients who experience a loss of treatment benefit after Week 52 if no scheduled visit available (SALT assessment required to confirm loss of treatment benefit) (see Sections 5.1.2 and 5.1.6). Not applicable to unscheduled visits conducted for other reasons.

[b] At Week 36, nonresponders (SALT >20) in the placebo treatment arm will be rescued to baricitinib (see Section 5.1.6 for details). For details on responder randomization at Week 52, refer to Section 5.1.6.

[c] A posterior–anterior view chest x-ray will be obtained locally at Screening (Visit 1), unless one has been performed in the past 6 months and the x-ray and/or the report is available.

[d] TB test(s) including PPD, QuantiFERON®-TB Gold, and T SPOT®. See Exclusion Criterion [27] for description of TB testing. In countries where the QuantiFERON-TB Gold test or T-SPOT is available, either test may be used instead of the PPD TB test. It is preferred that the QuantiFERON-TB Gold test

be performed centrally; the T-SPOT must be performed locally. Note: Patients with a history of active or latent TB who have documented evidence of appropriate treatment, have no history of re-exposure since their treatment was completed, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria are met. Such patients would not be required to undergo the protocol-specific TB testing, but must have a chest x-ray at screening.

[e] If PPD testing was chosen to test for TB, then the patient must return and have the PPD test read 48 to 72 hours after Visit 1 (post-PPD).

[f] The symptom-directed physical examination may be repeated at the investigator's discretion any time a patient presents with physical complaints.

[g] Presence of MPHL (AGA) will be assessed at screening using the Hamilton-Norwood Scales for male patients, through examination, and via patient interview; these will be re-evaluated at Weeks 36, 52, 104, and 200/ET. Presence of FPHL (AGA) will be assessed at Weeks 36, 52, 104, and 200/ET for female patients whose history of AGA is "no history" or unknown at screening.

[h] All sites/patients will obtain photographs of the scalp and eyebrows/eyelashes at Baseline. Photographs of the nails will be obtained only from patients with nail involvement (ClinRO ≥1) at Baseline. All sites/patients will obtain photographs of the scalp at Weeks 12, 36, and 52, and at Early Termination if it occurs before Visit 11 (Week 52). Photographs of the eyelash/eyebrows and/or nails will be captured at Weeks 12, 36, and 52 only from patients with AA involvement of those areas (ClinRO ≥1) at Baseline (Visit 2). See Section 9.1.4 for details.

[i] AA-IGA will be automatically derived from the SALT scores.

[j] A "Baseline/Screening" form is used at Visit 1. A "Since Last Visit" form used for all remaining visits. A suicidal ideation and behavior subscales excerpt is adapted for the assessment of 11 preferred ideation and behavior categories.

[k] The Self-Harm Follow-Up Form is only required if triggered by the Self-Harm Supplement Form

[l] Fasting lipid profile: patients should not eat or drink anything except water for 8 hours prior to sample collection. If a patient attends these visits in a nonfasting state, this will not be considered a protocol violation. Unscheduled lipid testing can be performed at the discretion of the investigator. For ET visits, collect fasting lipids when possible.

[m] Clinical chemistry will include the following value calculated from serum creatinine: eGFR (calculated using the CKD-EPI creatinine 2009 equation).

[n] For all women of childbearing potential, a serum pregnancy test (central laboratory) will be performed at Visit 1. Urine pregnancy tests (local laboratory) will be performed at Visit 2 and at all subsequent study visits. If required per local regulations and/or institutional guidelines, pregnancy testing can occur at other times during the study treatment period.

[o] For female patients ≥40 and <60 years of age who have had a cessation of menses for at least 12 months, an FSH test will be performed to confirm nonchildbearing potential (FSH ≥40 mIU/mL).

[p] For patients who are positive for HCV antibody, a follow-up test for HCV RNA will be performed automatically. Patients who are positive for HCV antibody and negative for HCV RNA may be enrolled.

[q] Patients who are positive for HBcAb and negative for HBV DNA may be enrolled. Any enrolled patient who is HBcAb positive, regardless of HBsAb status or level, must undergo HBV DNA testing per the schedule.

[r] Samples for the other clinical laboratory assessments must be drawn prior to receiving the first dose. PK samples are required at ET Visit only if the patient discontinues the study prior to Visit 6/Week 16. See Section 9.5 for PK details when IP should be taken in relationship to PK sampling and timing of PK sampling.

3. Introduction

3.1. Study Rationale

Although, in the past, alopecia areata (AA) has been considered by some to be a benign condition, severe AA is now recognized as a significant medical condition with emotional and psychosocial distress, including high prevalence of depression and anxiety (Colon et al. 1991; Hunt and McHale 2005; Villasante Fricke and Miteva 2015). Additionally, reduction in health-related quality of life experienced by patients with AA has been well documented (Jankovic et al. 2016; Liu et al. 2016; Rencz et al. 2016). Alopecia areata can be reversible – either spontaneously or after various types of treatments, but results are inconsistent (Islam et al. 2015) and there is currently no Food and Drug Administration- (FDA-) approved treatment for AA.

Baricitinib is an orally available, selective inhibitor of Janus kinases (JAKs). Janus kinases are a family of tyrosine kinases that mediate cytokine receptor signaling through phosphorylation and activation of signal transducers and activators of transcription (STAT) proteins. There are 4 known JAK family members: JAK1, JAK2, JAK3, and tyrosine kinase 2 (TYK2) (Clark et al. 2014). The relative affinity of JAK inhibitors for different members of the JAK kinase family allows for differentiation of JAK inhibitors in relation to their enzymatic inhibitory profile. In vitro assays indicate that baricitinib is a selective inhibitor of JAKs with potency and selectivity for JAK1 and JAK2 and less potency for JAK3 or TYK2 (Fridman et al. 2010).

Dual inhibition of JAK1 and JAK2, which may interrupt interferon gamma (IFNγ) signaling and other inflammatory pathways that contribute to the immunopathogenesis of AA, and clinical evidence with other JAK inhibitors support the investigation of baricitinib in the treatment of AA. This adaptive, operationally seamless, Phase 2/3, placebo-controlled study is designed to select up to 2 doses of baricitinib and assess their efficacy and safety for the treatment of severe (Severity of Alopecia Tool [SALT] score of 50% to 94%) and very severe AA (SALT score of 95% to 100%). Additionally, it will include a randomized withdrawal to explore the persistence of treatment effects and efficacy of retreatment upon relapse. This operationally seamless design will provide longer-term safety data earlier in the clinical development program, compared to running separate Phase 2 and Phase 3 studies.

Lilly has initiated a clinical program to evaluate the efficacy and safety of baricitinib in adult patients with severe (SALT score of 50% to 94%) and very severe AA (SALT score of 95% to 100%) AA of the scalp.

This clinical program currently comprises 2 trials:

- Study JAHO described in this protocol, is a double-blind, placebo-controlled adaptive Phase 2b/3 trial. In the Phase 2 dose-ranging portion of the study, 3 dose regimens were tested: baricitinib 1-mg QD, baricitinib 2-mg QD, and baricitinib 4-mg QD. The results of the interim analysis of the Phase 2 portion confirmed the potential of baricitinib to restore hair growth and supported further evaluation of the 2-mg QD and 4-mg QD regimen in the Phase 3 portion of the study. See Section 5.5 (Justification for Dose) for more information.

- Study I4V-MC-JAIR (JAIR), is also a double-blind, placebo-controlled Phase 3 trial and will evaluate the same doses.

Studies JAHO and JAIR have identical inclusion and exclusion criteria and have the same design and primary and key secondary objectives for the 36-week, double-blinded, placebo-controlled Treatment Period (Period 2).

3.2. Background

Alopecia areata is an autoimmune disease usually characterized by patches of nonscarring hair loss that can affect the scalp, face, and body, with a lifetime prevalence of approximately 2% (Wasserman et al. 2007; Islam et al. 2015; Korta et al. 2018). The extent of hair loss in this condition can vary, including total hair loss on the scalp and loss of facial and/or other body hair (Islam et al. 2015). Alopecia areata can affect children, adolescents, and adults (Hordinsky and Donati 2014).

Alopecia areata is one of the more common autoimmune diseases and its visible phenotypic characteristic hair loss and lack of FDA-approved treatments provide a unique opportunity for novel treatment approaches with a drug with immune modulatory effects such as baricitinib. Clinical manifestations of AA (hair loss and nail effects) may be constant or sporadic and can affect different areas of the scalp, face, or body at different times during life, or cause hair loss in the entire scalp or the entire body at once (Olsen 2011). As an autoimmune disease in which the patient's own T cells produce cytokines that inhibit hair follicle growth, AA may present as patchy AA, with partial hair loss in one or multiple well-circumscribed round patches in the scalp (or other hair-bearing site), or it can progress to hair loss involving the entire scalp (alopecia totalis [AT]), or extend to complete hair loss of head and body (alopecia universalis [AU]) (Olsen 1999; Olsen et al. 2004). Patchy alopecia, AT, and AU are considered part of the clinical spectrum of the same disease (Hordinsky and Junqueira 2015). Alopecia areata can cause significant emotional and psychosocial distress (Hunt and McHale 2005). A 66% to 74% lifetime prevalence of psychiatric disorders has been reported in patients with AA, with a 38% to 39% lifetime prevalence of depression and a 39% to 62% prevalence of generalized anxiety disorder (Colon et al. 1991; Villasante Fricke and Miteva 2015). Quality of life is also consistently diminished in patients with AA (Jankovic et al. 2016; Liu et al. 2016; Rencz et al. 2016). The Global Burden of Disease Study (GBD) determined the burden of AA to be a mean of 19.4 years of healthy life lost (Karimkhani et al. 2015). However, this GBD study did not account for other potential harms, such as emotional distress and financial impact (Karimkhani et al. 2015).

Insights into the immunopathogenesis of AA began with the recognition of the hair follicle as being an immune-privileged site like the eye and testes (Paus et al. 2005). Next, disruption of this immune privilege occurs upon follicular influx by auto-reactive CD8+ T cells, leading to increases in major histocompatibility complex (MHC) class I and II antigens and inflammation disrupting hair follicle biology (Islam et al. 2015; Strazzulla et al. 2018). Activation of the pathogenic T cells leads to IFNγ production which contributes both to enhanced MHC class I and II antigens and interleukin-15 (IL-15) (Islam et al. 2015; Strazzulla et al. 2018) accompanied by additional cytokines including

IL-2, IL-13, IL-23, and thymic stromal lymphopoietin (Suárez-Fariñas et al. 2015). All of these inflammatory-related cytokines are dependent on JAK/STAT signaling, and of particular note IFNγ utilizes JAK1 and JAK2. Animal models support the theory of AA in which autoreactive T cells (NKG2D+) drive the hair loss by increasing IFNγ and inflammatory gene expression signatures as noted above, which could be reversed using JAK inhibition in mice (Xing et al. 2014).

There are currently no FDA-approved treatments for AA. While not all patients with AA are likely to require treatment because AA is often mild and intermittent (Tan et al. 2002; Wasserman et al. 2007), for patients with severe and persistent disease, safe and efficacious treatments are lacking. Use of a JAK 1/JAK 2 inhibitor such as baricitinib may be able to reinitiate production of mature terminally differentiated follicles at sites of prior inflammation.

3.3. Benefit/Risk Assessment

Several retrospective case series, case reports, and open-label studies have been published describing positive outcomes in patients with severe (SALT score of 50% to 94%) or very severe AA (SALT score of 95% to 100%), treated with different JAK inhibitors for 4 to 18 months. In these reports, more than one-half of patients experienced hair regrowth of 50% or more after beginning treatment with a JAK inhibitor (Xing et al. 2014; Mackay-Wiggan et al. 2016; Craiglow et al. 2017; Liu et al. 2017; Jabbari et al. 2018; Liu et al. 2018;), including baricitinib (Jabbari et al. 2015). A 66-patient, open-label study reported that approximately one-third of patients achieved hair regrowth of >50% after treatment with tofacitinib; however, these patients were treated with a JAK inhibitor for only 3 months (Kennedy Crispin et al. 2016). Results from the interim analysis of the Phase 2 portion of Study JAHO showed that the 2-mg and 4-mg baricitinib doses demonstrated numerical superiority over the 1-mg dose and placebo, in hair regrowth, after 12 ($SALT_{30}$) and 16 ($SALT_{50}$) weeks of treatment with no new safety signal. These results have been confirmed by a second interim analysis of Phase 2 portion of Study JAHO where the 2-mg dose and 4-mg dose were statistically significantly superior to placebo on the primary endpoint (SALT≤20) after 36 weeks of treatment.

Serious infections, venous thromboembolic events (VTEs), hepatotoxicity, and fetal malformations were identified as important potential risks with baricitinib in RA studies Although infections were seen in approximately one-half of the study population exposed to baricitinib in the RA program, only 3.6% of patients reported a serious treatment-emergent infection, and rates were similar in both baricitinib- and placebo-treated patients. In the Phase 2 atopic dermatitis (AD) study, approximately 25% of patients treated with the 4-mg dose experienced a treatment-emergent adverse events (TEAE) of infection and infestation, compared to 20% of patients treated with placebo. There were no serious infections, opportunistic infections, or herpes zoster infections reported during the treatment period in the AD Phase 2 study. There was no serious event reported at the time of the second interim analysis of the Phase 2 portion of Study JAHO after 36 weeks of treatment. The nonserious infections for baricitinib noted in the RA program (upper respiratory tract infections, herpes zoster, herpes simplex) are readily diagnosed and manageable, and typically resolve without long-term sequelae. It is recommended that, where indicated, herpes zoster vaccination will be offered to

patients before receiving baricitinib; herpes zoster infections, when they have occurred, have generally been mild-to-moderate, localized, and without long-term problems. Exclusion criteria have been added to the protocol to limit enrollment of patients who are at increased risk for infection.

Cases of hepatotoxicity have not been identified with baricitinib use, but increases in alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin (TBL) have been seen in RA patients. Most of these increases improved with continued use or temporary discontinuation of baricitinib with no long-term effects. Exclusion criteria to not enroll patients with liver failure or increased liver analytes, appropriate monitoring of hepatic analytes, and discontinuation criteria have been included in the protocol.

In toxicology studies, fetal malformations were reported at doses higher than what is used in human patients. Only a small number of patients have become pregnant in baricitinib clinical trials, and there have been no reports of fetal malformations in these pregnancies. In current study protocols, pregnant patients are excluded from entry, contraceptive use is defined in the inclusion criteria, and patients who become pregnant are discontinued from the trial.

Venous thromboembolic events have been determined to be an important potential risk for baricitinib. There was a numerical imbalance in reports of VTEs in the 24-week, placebo-controlled period of the Phase 3 trials of patients with RA. Available evidence does not establish a causal association. The exposure-adjusted incidence rate of VTE for baricitinib-treated RA patients over long-term exposures was similar to the background rates published in the literature for the target population (i.e., RA). There was no pattern of increased or decreased risk during long-term exposures, and cases observed with baricitinib were confounded by 1 or more recognized risk factors for a VTE (eg, history of VTE, increased body mass index [BMI], older age). Venous thromboembolic event risk can be managed through risk mitigation strategies. Exclusion and discontinuation criteria have been added to the protocol to limit participation of patients who are at increased risk for VTE. Patients will be excluded if they have a previous history of a VTE, or have a combination of 2 or more risk factors for a VTE (eg, increased BMI, older age, combination use of oral contraceptive with smoking). History of a previous VTE is the strongest risk factor for development of a second VTE, and the existing exclusion criteria for patients with previous history of a VTE will also serve to exclude patients with other risk factors.

Presymptomatic screening for hypercoagulable conditions is not currently recommended, even in the prescribing information for medications with known increased risk for VTEs. Although temporary interruption of investigational product (IP) is not a requirement at times of increased potential risk for VTE (eg, surgery, significant air travel, or other situations involving prolonged immobilization), we recommend following appropriate VTE prophylaxis guidelines to help manage the VTE risk under these circumstances.

There was no VTE reported at the time of the second interim analysis of the Phase 2 portion of Study JAHO after 36 weeks of treatment.

Therefore, in the context of the cumulative knowledge, the benefit/risk balance for baricitinib for the treatment of adult patients with severe AA is assessed to be favorable.

More detailed information about the known and expected benefits, risks, serious adverse events (SAEs), and reasonably anticipated adverse events (AEs) of baricitinib are to be found in the Investigator's Brochure (IB).

4. Objectives and Endpoints

Table JAHO.2 shows the objectives and endpoints of the study.

Table JAHO.2. Objectives

Objectives	Endpoints
Primary Objective	
To test the hypothesis that the 4-mg dose or 2-mg dose of baricitinib is superior to placebo in the treatment of patients with severe or very severe AA	• Proportion of patients achieving SALT ≤20 at Week 36.
Key Secondary Objectives (Double–Blind, Placebo-Controlled Treatment Period) *These are pre-specified objectives that will be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as measured by **physician-assessed** signs and symptoms of AA	• Proportion of patients achieving SALT ≤20 at Weeks 16 and 24. • Percent change from baseline in SALT score at Week 36. • Proportion of patients achieving a $SALT_{50}$ at Week 12. • Proportion of patients achieving a $SALT_{90}$ at Week 36. • Proportion of patients achieving an absolute SALT ≤10 at Weeks 24 and 36. • Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week 36 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline). • Proportion of patients achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week 36 (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline).
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as assessed by a **PRO measure**	• Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement from Baseline at Week 36 among patients with a score of ≥3 at Baseline.
Other Secondary Objectives (Double–Blind, Placebo-Controlled Treatment Period) *These are pre-specified objectives that will NOT be adjusted for multiplicity.*	
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as measured by **physician-assessed** signs and symptoms of AA	• Proportion of patients achieving $SALT_{50}$ at Weeks 16, 24, and 36. • Proportion of patients achieving a $SALT_{75}$ at Weeks 24 and 36. • Proportion of patients achieving a $SALT_{90}$ at Week 24. • Change from Baseline in SALT score at Weeks 12, 16, 24, and 36. • Percent change from Baseline in SALT score at Weeks 12, 16, and 24. • Time to achieve SALT ≤20.

Objectives	Endpoints
	• Proportion of patients achieving $SALT_{100}$ at Weeks 24 and 36. • Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Weeks 16 and 24 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline). • Proportion of patients achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from baseline at Weeks 16 and 24 (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline).
To compare the efficacy of baricitinib 4-mg or 2-mg to placebo in AA during the double-blind, placebo-controlled treatment period as assessed by **PRO measures** and quality of life tools	• Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement from Baseline at Weeks 12 and 24 among patients with a score of ≥3 at Baseline • Proportion of patients achieving PRO Measure for EB 0 or 1 with ≥2-point improvement from baseline at Weeks 16, 24, and 36 (among patients with PRO Measure for EB ≥2 at Baseline). • Proportion of patients achieving PRO Measure for EL 0 or 1 with ≥2-point improvement from baseline at Weeks 16, 24, and 36 (among patients with PRO Measure for EL ≥2 at Baseline) • Mean change from Baseline in HADS-A and HADS-D total scores at Weeks 24 and 36
Other Secondary (Patients entering Randomized Withdrawal) *These are pre-specified objectives that will NOT be adjusted for multiplicity.*	
To compare the maintenance of efficacy for patients randomized to remain on baricitinib, compared with patients randomized to placebo at Week 52 of the long-term extension period, as measured by **physician-assessed** signs of AA	• Proportion of patients maintaining SALT ≤20 at Weeks 64, 76, 88, 104, 120, 136, 152, 168, 184, and 200. • Proportion of patients experiencing a loss of treatment benefit (>20-point absolute worsening in SALT score) at Weeks 64, 76, 88, 104, 120, 136, 152, 168, 184, and 200. • Time to loss of treatment benefit (>20-point absolute worsening in SALT score).
For patients experiencing loss of treatment benefit after randomization to placebo at Week 52: • To evaluate the recapture of efficacy for patients who were retreated after experiencing a loss of treatment benefit during the long-term maintenance period as measured by **physician-assessed** signs of AA • To evaluate the recapture of efficacy for patients who were retreated after experiencing a loss of treatment benefit during the long-term maintenance period as assessed by **PRO** and quality of life tools	• Proportion of patients that achieve a SALT ≤20 at 12, 16, 24, and 36 weeks of retreatment with baricitinib. • Percent change in SALT score at 12, 16, 24, and 36 weeks of retreatment with baricitinib. • Proportion of patients with a PRO for Scalp Hair Assessment score of 0 or 1 at 12, 16, 24, and 36 weeks of retreatment with baricitinib.

Objectives	Endpoints
Exploratory Objectives may include evaluating the response to baricitinib treatment regimens on clinical measures and patient-reported outcomes. These endpoints may include dichotomous endpoints or change from baseline for the following measures: SALT, AA-IGA, $SALT_{30,}$ ClinROs for Nail Appearance, Eyebrows, and/or Eyelash Hair Loss, PROs for Scalp Hair Assessment, Eyebrows, and Eyelashes, Nail Appearance and Eye Irritation, Skindex-16 AA, SF-36, EQ-5D-5L, HADS. Assessments of efficacy may be performed beyond Week 104 up to Week 200. In addition, baricitinib pharmacokinetics will be characterized in the AA population and relationships between exposure and study endpoints will be explored.	

Abbreviations: AA = alopecia areata; AA-IGA = Alopecia Areata Investigator Global Assessment; ClinRO = clinician-reported outcome; EB = eyebrow; EL = eyelash; EQ-5D-5L= European Quality of Life – 5 Dimensions – 5 Level; HADS = Hospital Anxiety and Depression Scale; PRO = patient-reported outcome; SALT = Severity of Alopecia Tool; $SALT_{30}$ = at least 30% improvement from Baseline in SALT score; $SALT_{50}$ = at least 50% improvement from Baseline in SALT score; $SALT_{75}$ = at least 75% improvement from Baseline in SALT score; $SALT_{90}$ = at least 90% improvement from Baseline in SALT score; $SALT_{100}$ = 100% improvement from Baseline in SALT score; SF-36 = Short Form-36 Health Survey acute version 2; Skindex-16 AA = Skindex-16 Adapted for Alopecia Areata.

5. Study Design

5.1. Overall Design

Study I4V-MC-JAHO (JAHO) is an adaptive, operationally seamless, Phase 2/3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study designed to identify up to 2 doses of baricitinib to be evaluated further in the Phase 3 portion of the study. The 2-mg and 4-mg doses of baricitinib were selected as a result of the Phase 2 interim analysis; therefore, efficacy and safety of baricitinib 2-mg and 4-mg will be compared to placebo in adult patients with severe (SALT score of 50% to 94%) or very severe (SALT score of 95% to 100%) scalp AA during the Phase 3 portion of Study JAHO. Approximately 725 adult patients will be enrolled into Study JAHO. Approximately 100 patients will be enrolled into the Phase 2 portion of the study and approximately 625 patients will be enrolled into the Phase 3 portion of the study.

Patients must have a current AA episode of more than 6 months' duration prior to screening (Visit 1), with at least 50% scalp involvement at screening AND Baseline (Visits 1 and 2) with no spontaneous improvement (no more than a 10 point reduction in SALT) over the past 6 months. Patients with a current episode of severe or very severe AA of more than 8 years will not be eligible for inclusion in the study unless episodes of regrowth, spontaneous or under treatment, have been observed on the affected areas of the scalp over the past 8 years.

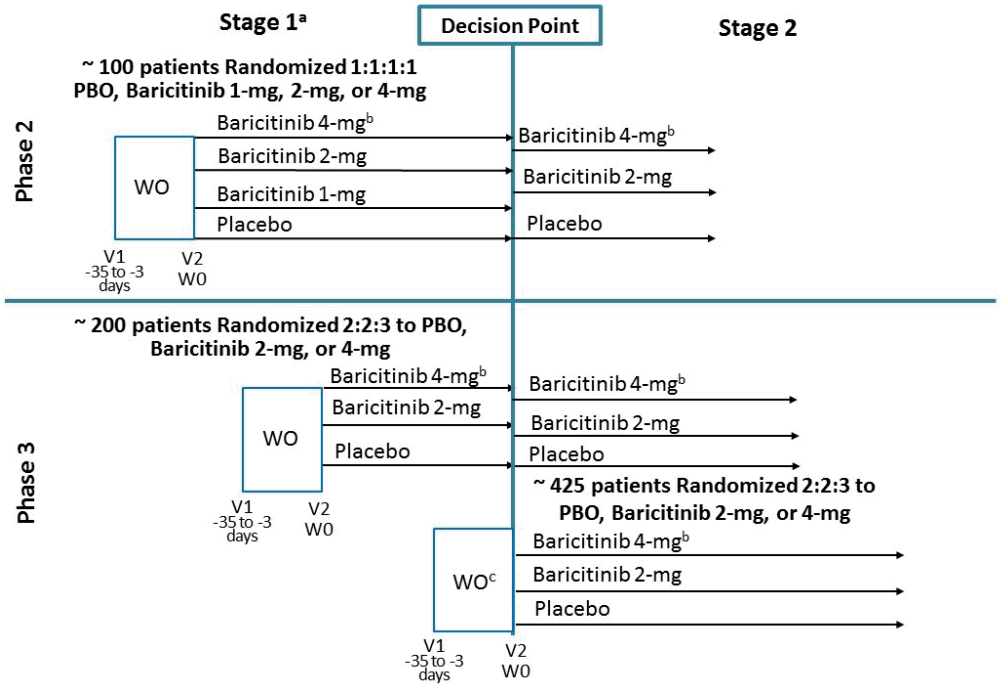
All procedures to be conducted during the study, including timing of all procedures, are indicated in the Schedule of Activities (SOA) (Section 2). Section 9.4.4 describes collection of laboratory samples; Appendix 2, Appendix 4, and Appendix 5 list the specific laboratory tests that will be performed for this study. Study governance considerations are described in detail in Appendix 3. Section 10.3.7 outlines information regarding the interim analyses and Decision Point Committee.

Study Stages

The enrollment of patients in the study will be divided into 2 stages, which are separated by the Decision Point. Different randomization schemes at Baseline (Visit 2) will be used by the interactive web-response system (IWRS) (during Stage 1 [Phase 2], Stage 1 [Phase 3] and Stage 2 [see Figure JAHO.1]).

- Stage 1: the time from study start until the Decision Point. A maximum of approximately 300 patients will be randomized during Stage 1, before the Decision Point. See Section 5.1.1 for additional details.
- Decision Point: the point in time when up to 2 baricitinib doses will be selected to continue in Stage 2 or the study will be stopped for futility based on the outcome of the interim analysis. See Section 5.1.2 for additional detail.
- Stage 2: the time after the Decision Point until the end of the study during which the remaining patients (approximately 425 patients) will be enrolled into the Phase 3 portion of the study. See Section 5.1.3 for additional detail.

- Based on the Week 12 interim analysis of the Phase 2 population of this study, baricitinib 2-mg QD (low dose) and baricitinib 4-mg QD (high dose) were selected at Decision Point to continue into Stage 2 of Study JAHO.



Abbreviations: eGFR = estimated glomerular filtration rate; PBO = placebo; QD = once daily administration; V = visit; WO = washout.

a Patients randomized during Stage 1 who are in the treatment arm that is discontinued (baricitinib 1-mg) will be transitioned to the 4-mg dose of baricitinib remaining in the study after the Decision Point.

b The maximal baricitinib dose for patients with renal impairment (defined as eGFR <60 mL/min/1.73 m^2) will be 2-mg QD (see Section 5.5.1).

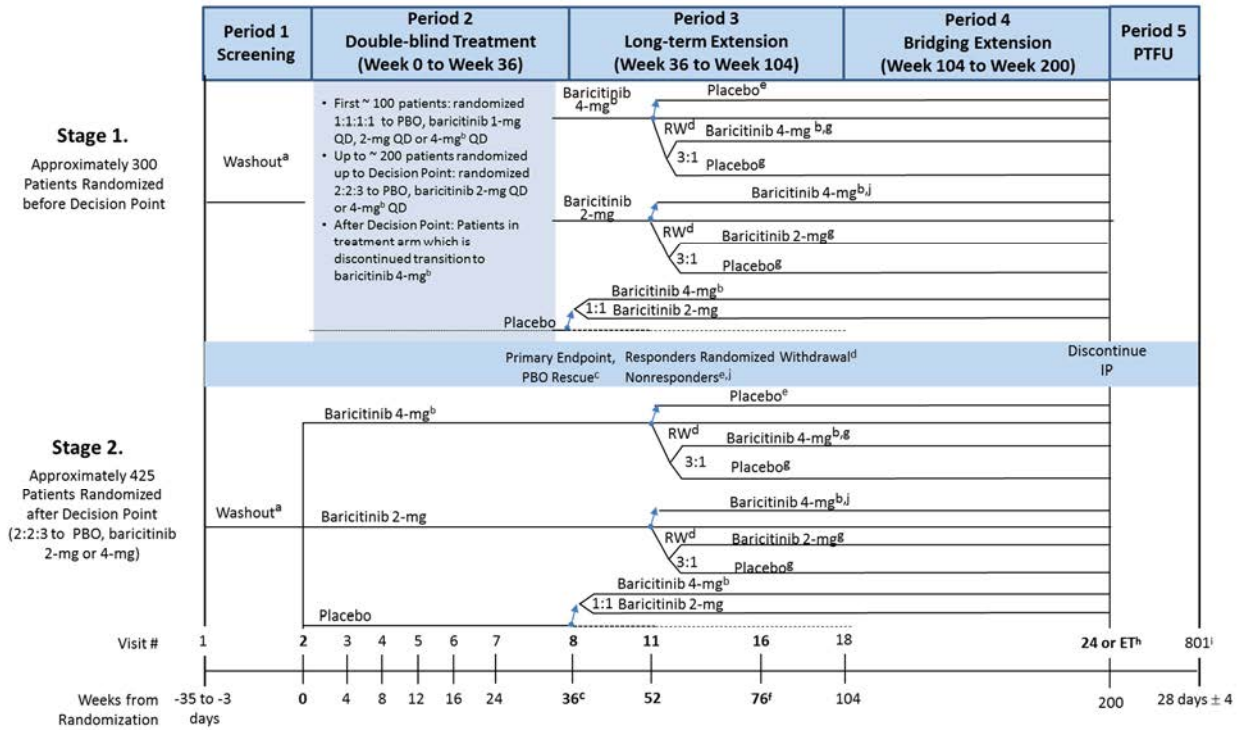
c Some of the first patients randomized during Stage 2 may have begun washout during Stage 1.

Figure JAHO.1. Illustration of randomization schemes and enrollment during the 2 stages of the protocol, before and after the Decision Point.

Study Design

The study design includes 5 periods: a 5-week screening period, a 36-week double-blind, placebo-controlled treatment period, a 68-week long-term extension period, a 96-week bridging extension, and a posttreatment follow-up period. The SOA (Section 2) will be the same for patients randomized during Stage 1 and Stage 2 except some patients enrolled during Stage 1 may have an unscheduled visit(s) after the Decision Point (see Section 5.1.6).

Figure JAHO.2 illustrates the study design. The full visit schedule is outlined in the SOA (Section 2).



Abbreviations: AA-IGA = Alopecia Areata Investigator Global Assessment; ClinRO = clinician reported outcome; EC = exclusion criterion; eGFR = estimated glomerular filtration rate; ET = early termination; IP = investigational product; PBO = placebo; PTFU = posttreatment follow-up; QD = once daily administration; RW = randomized withdrawal; SALT = Severity of Alopecia Tool; V = Visit.

[a] Applicable to all patients at time of screening. See EC [9] for treatments that will require washout.

[b] The maximal baricitinib dose for patients with renal impairment (defined as eGFR <60 mL/min/1.73 m^2) will be 2-mg QD (see Section 5.5.1).

[c] At Week 36 patients in the placebo treatment arm who have **NOT** achieved SALT ≤20 will be rescued and re-randomized in a 1:1 ratio to baricitinib 2-mg or baricitinib 4-mg. All patients in the baricitinib treatment arms will continue in their current treatment arm regardless of treatment response at Week 36. Patients in the placebo arm who have achieved a SALT ≤20 will remain on placebo at Week 36. These patients who have experienced spontaneous regrowth on the scalp will remain on placebo for the remainder of the trial, even if relapse is observed later during the study.

[d] At Week 52, responders, SALT ≤20, who are eligible (i.e., stayed on the same dose of baricitinib from initial randomization at Visit 2) will be randomized in a 3:1 ratio to either stay on their current dose of baricitinib or transition to placebo (randomized withdrawal).

[e] Patients who have been in the baricitinib 4-mg treatment group from Baseline and who have never achieved a SALT ≤20 by Week 52 AND do not have a ≥2 point improvement in ClinRO measure for eyebrow or eyelash hair loss (nonresponders) at Week 52 will be automatically transitioned to placebo. See footnote "f" for discontinuation criteria at Week 76.

[f] Patients who are nonresponders, SALT >20, at Weeks 52 AND 76 will be automatically discontinued from the study at Week 76, unless they have a ≥2-point improvement from baseline in ClinRO measure for eyebrow or eyelash hair loss. See Section 5.1.6 for more details.

[g] Responders who experience a loss of treatment benefit after Week 52 (>20-point absolute worsening in SALT score) who were randomized to placebo at Week 52 (randomized withdrawal) will be retreated with their baricitinib dose, as randomized at Baseline (Visit 2). Patients who were randomized to remain on baricitinib (randomized withdrawal) will continue to receive the same dose of baricitinib. See Section 5.1.6 for more detail.

[h] ET Visit is required for patients that terminate IP early. Patients who remain in the study for more than 28 days after discontinuation of IP do not need a separate follow-up visit (V801).

[i] V801 occurs approximately 28 days after the last dose of IP. Patients who have completed Week 200 and will continue on marketed product beyond Week 200 do not need to complete Period 5.

[j] Patients who are nonresponders at Week 52 and who have been in the baricitinib 2-mg treatment group from Baseline will be rescued to baricitinib 4-mg.

Figure JAHO.2. Illustration of study design for Clinical Protocol I4V-MC-JAHO.

5.1.1. Stage 1 (Includes Phase 2 Portion and Phase 3 Portion before Decision Point)

The objective in the Phase 2 portion of JAHO is to identify up to 2 baricitinib doses that will continue into Phase 3, or to stop the study early for futility.

Stage 1:

A maximum of approximately 300 patients are expected to enroll during Stage 1.

- The first approximately 100 patients (Phase 2) will be randomized in a 1:1:1:1 ratio to placebo, baricitinib 1-mg once daily (QD), baricitinib 2-mg QD, or baricitinib 4-mg QD (see Figure JAHO.1). Patients will be stratified at randomization by geographic region (North America and Japan), and duration of current episode at Baseline (less than 4 years versus at least 4 years).
- An interim analysis, as described in Section 10.3.7, will be conducted after 100 patients who have been randomized and treated (approximately 25 from each dose level) have reached Week 12 (Visit 5) or discontinued early, to determine which of the 3 baricitinib doses qualify to continue into Phase 3.
- Patients will continue to be enrolled in Stage 1 until the Decision Point is reached: approximately up to 200 additional patients will be included in the Phase 3 portion of the study and will be randomized in a 2:2:3 ratio to placebo, baricitinib 2-mg QD or baricitinib 4-mg QD, as it is anticipated that the 1-mg dose will be the dose most likely to be dropped after the Decision Point. Patients will be stratified at randomization by geographic region (North America, Asia, and Rest of World) and duration of current episode at Baseline (less than 4 years versus at least 4 years).

Note: Patients with an estimated glomerular filtration rate (eGFR) of <60 mL/min/1.73 m^2 who are randomized to baricitinib 4-mg will receive a dose of 2-mg QD.

Patients enrolled during Stage 1 will undergo the same visits and procedures described below for patients enrolled during Stage 2, as shown in Figure JAHO.2; however, some patients enrolled during Stage 1 may have up to 2 unscheduled visits after the Decision Point, as described in Section 5.1.2.

5.1.2. Decision Point

The Decision Point is the point in time when up to 2 baricitinib doses will be selected to continue in Stage 2, or the study will be stopped for futility based on the outcome of the Phase 2 population, Week 12 interim analysis. Based on this interim analysis, the 4-mg and 2-mg doses of baricitinib were selected to continue into the Phase 3 portion of Study JAHO.

Patients in the baricitinib treatment arm that is not selected at the Decision Point (baricitinib 1-mg) will remain in the study, and they will be transitioned to the highest dose of baricitinib remaining in the study (4-mg) at the next study visit following the Decision Point.

Patients and sites will remain blinded to treatment allocation after the Decision Point and, therefore, will not know which patients will be transitioned. Transition will automatically occur at the next visit after Decision Point; this will be referred to as the Transition Visit. A patient should be seen within 8 weeks following the Decision Point being communicated to the sites. If there is not a regularly scheduled visit during this timeframe, patients may be brought in for an unscheduled visit. A patient should be seen within 8 weeks following the Transition Visit to perform laboratory tests for safety review. If there is not a regularly scheduled visit during this timeframe, patients may be brought in for an unscheduled visit (see Schedule of Activities [Section 2] for detailed requirements). After the Decision Point, all patients enrolled during Stage 1 will follow all protocol procedures for Periods 2, 3, and 4.

5.1.3. Stage 2 (Remaining Portion of Phase 3 Study after the Decision Point)

Approximately 425 additional patients will be enrolled in a 2:2:3 ratio to placebo, baricitinib 2-mg QD, or baricitinib 4-mg QD, as shown in Figure JAHO.2. Patients will be stratified at randomization by geographic region (North America, Asia, and Rest of World) and duration of current episode at Baseline (less than 4 years versus at least 4 years).

5.1.4. Period 1: Screening (-35 to -3 days)

The duration of the Screening Period is between 3 and 35 days prior to Visit 2 (Week 0). Therapies that must be washed out or discontinued before randomization (Visit 2) are listed in exclusion criteria [9]. Patients who receive a purified protein derivative (PPD) skin test at Visit 1 will return 48 to 72 hours later to read the skin test. Patients who are eligible for herpes zoster vaccine (per local guidelines) and have not previously received the vaccine will be encouraged to do so prior to randomization. Vaccination with the live herpes zoster vaccination (single injection) must occur at least 28 days (4 weeks) prior to randomization. Vaccination with the non-live herpes zoster vaccine requires 2 injections administered at least 8 weeks apart. It is recommended that patients who have initiated vaccination with the non-live herpes zoster vaccine receive the second dose at least 4 weeks prior to randomization. For both vaccines, completion of the vaccination may require rescreening the patients to comply with the maximum allowed duration of the screening period. In addition, investigators should review the vaccination status of their patients and follow the local guidelines for vaccination of those ≥18 years of age with non-live vaccines intended to prevent infectious disease prior to entering patients into the study. Patients who meet all of the inclusion criteria (Section 6.1) and none of the exclusion criteria (Section 6.2) will continue to Visit 2.

5.1.5. Period 2: Double-Blind Treatment Period (Weeks 0 – 36)

At Visit 2 (Week 0, Baseline), study eligibility for each patient will be reviewed on the basis of all inclusion (Section 6.1) and exclusion (Section 6.2) criteria and laboratory test results. Patients who meet all eligibility criteria will proceed to randomization and begin the 36-week double-blind, placebo-controlled treatment period (Study Period 2).

At Visit 2, after laboratory samples are collected and all assessments are completed, patients will take the first dose of IP in the clinic. Pharmacokinetic (PK) samples will be drawn 15 minutes and 1 hour after the dose. See Section 9.5 for the timing of PK sampling at Visits 3, 4, 5, and 6 or the early termination (ET) (if ET occurs prior to Visit 6).

Please refer to Section 5.1.1 for enrollment and randomization details of Stage 1 (before Decision Point), Section 5.1.2 for details of the Decision Point, and Section 5.1.3 for enrollment and randomization details of Stage 2.

During Study Period 2 (double-blind treatment), medications listed in Section 7.7.2 will be prohibited.

At Baseline (Visit 2), photographs of the scalp (4 planes), 1 photograph (frontal view) of the face and scalp, and photographs of the eyebrows/eyelashes will be obtained from all patients; photographs of nails will be obtained only from patients who have nail involvement (clinician reported outcome [ClinRO™ Measure for Nail Appearance] ≥1) at Baseline. At Week 12 (Visit 5), photographs of the scalp (4 planes) and 1 photograph (frontal view) of the face and scalp will be obtained from all patients and photographs of eyelashes/eyebrows and/or nails will be captured only from patients who have AA involvement (ClinRO ≥1) of those areas at Baseline, as assessed by investigator (see SOA [Section 2] and Section 9.1.4). Missed photographs will not be considered a protocol violation. Detailed requirements for photography are located in the Photography Manual.

The primary efficacy endpoint will be at Week 36 (Visit 8). If a patient discontinues IP prior to Week 36 (Visit 8), the patient is encouraged to remain in the study and follow the regular visit schedule to provide the primary efficacy and safety data through Week 36 (Visit 8).

5.1.6. Period 3: Long-Term Extension

All patients who have completed Study Period 2 (Week 36) will enter the long-term extension period (up to 68-weeks of additional treatment [Study Period 3]). Patients will continue on their current treatment assignment unless predefined criteria are met for rescue.

Photographs will be obtained at Weeks 36 and 52, as described in the SOA (Section 2) and Section 9.1.4; however, missed photographs will not be considered a protocol violation. Photographs will be obtained at ET for patients who discontinue prior to Week 52.

During Study Period 3 (long-term extension), medications listed in Section 7.7.2 will be prohibited.

At Week 36

The predicted enrollment rate for Study JAHO suggests that some of the patients enrolled during Stage 1 may reach Week 36 before the Decision Point is achieved. See Section 5.1.2 for additional details on transitioning patients after the Decision Point:

- Patients in the placebo treatment arm who have not achieved SALT ≤20 at Week 36, AND who reach the Week 36 Visit:
    - prior to the Decision Point, will be rescued to baricitinib and randomized in a 1:1 ratio to baricitinib 2-mg or baricitinib 4-mg.
    - after the Decision Point, will be rescued to baricitinib and randomized in a 1:1 ratio to baricitinib 2-mg or baricitinib 4-mg.
- Patients in the placebo arm who have achieved a SALT ≤20 will remain on placebo at Week 36. These patients who have experienced spontaneous regrowth will remain on placebo for the remainder of the trial, even if relapse is observed later during the study.
- All patients in the baricitinib treatment arms will continue in their current treatment group, regardless of their treatment response at Week 36 (refer to Week 52 and Week 76 sections below for additional rescue information and discontinuation criteria).

At Week 52

Responders: SALT ≤20

- Patients in baricitinib treatment arms who achieve a SALT ≤20 at Week 52 (responders) are eligible for randomized withdrawal, provided that they have stayed on the same dose of baricitinib from initial randomization (Visit 2).
- Responders who have had a change in dose after the Decision Point, or patients who were rescued to baricitinib at Week 36, will not be eligible for randomized withdrawal and will remain in their same treatment group.
- Eligible patients will be automatically randomized in a blinded manner by the IWRS in a 3:1 ratio to either stay on their current dose of baricitinib or transition to placebo (randomized withdrawal).
- Any patients in the placebo treatment arm at Week 52 who have achieved a SALT ≤20 will remain on placebo.
- Responders who experience a loss of treatment benefit after Week 52 (defined as >20-point absolute worsening in total SALT score), and who:
    - were randomized to placebo at Week 52 (randomized withdrawal), will be automatically retreated with their baricitinib dose, as randomized at Baseline (Visit 2).
    - remained on baricitinib at Week 52 (randomized withdrawal), will continue to receive the same dose of baricitinib.
    - were randomized to placebo at Baseline (Visit 2), will remain on placebo.

Note: An unscheduled visit may be used after Week 52 to assess the patient who reports a loss of treatment response (see Schedule of Activities [Section 2] for detailed requirements).

Nonresponders: SALT >20

- Patients who have been in the baricitinib 4-mg treatment group from Baseline AND have never achieved a SALT ≤20 by Week 52 AND do not have a ≥2-point improvement from Baseline in ClinRO measure for eyebrow or eyelash hair loss at Week 52 will be automatically transitioned to placebo.
- Patients who have been in the baricitinib 4-mg treatment group and have achieved a SALT ≤20 before Week 52 and have lost response, will remain on baricitinib 4-mg dose.
- Those who have been in the baricitinib 2-mg treatment group from Baseline will be rescued to baricitinib 4-mg.
- Those who were transitioned to baricitinib 4-mg after the Decision Point, or were rescued to baricitinib at Week 36, will continue in their current treatment arm at Week 52.

Week 76

- Patients who are nonresponders (SALT >20) at Weeks 52 AND 76 will be automatically discontinued from the study at Week 76, unless they have a ≥2-point improvement from baseline in ClinRO measure for eyebrow or eyelash hair loss.

5.1.7. Period 4: Bridging Extension

Patients who have completed Week 104 and have not met criteria for permanent discontinuation will have the possibility to remain in the trial for up to 96 additional weeks (up to Week 200).

During Period 4, patients will continue to receive the same treatment they received during Period 3.

Responders who had been randomized to placebo at Week 52 (randomized withdrawal) and have remained on placebo, will have the possibility to be retreated with their baricitinib dose as randomized at Baseline (Visit 2) if they experience a loss of benefit (defined as >20-point absolute worsening in total SALT score).

5.1.8. Period 5: Posttreatment Follow-Up

Patients who complete the study through Visit 24 (Week 200), will have a posttreatment follow-up visit (Visit 801) approximately 28 days after the last dose of IP. Patients who have completed Week 200 and who will continue on marketed product beyond Week 200 do not need to complete Period 5 (Visit 801).

Patients who have received at least 1 dose of IP and discontinue early from the study must have an ET, and return for the posttreatment safety follow-up visit (Visit 801) approximately 28 days after the last dose of IP.

Patients who have discontinued IP product but remain in the study for more than 28 days without IP will have an ET if they choose to discontinue early; however, a separate follow-up visit (V801) is not required.

5.2. Number of Participants

Planned enrollment is approximately 100 patients in Phase 2 and approximately 625 patients in Phase 3. A maximum of approximately 300 patients will be enrolled during Stage 1 (approximately 100 patients in Phase 2, and 200 patients in Phase 3). The remaining approximately 425 patients for Phase 3 will be enrolled during Stage 2. Approximately 1035 patients in total will be screened to meet the planned total enrollment of 725 patients.

5.3. End of Study Definition

End of study is the date of the last visit or last scheduled procedure shown in the SOA (Section 2) for the last patient.

5.4. Scientific Rationale for Study Design

Available data strongly suggest that AA results from an immune-mediated perturbation of hair follicle biology due to pro-inflammatory mediators, such as IFNγ and other cytokines utilizing JAK-STAT signaling. Therefore, there is strong scientific and immunological rationale for studying the ability of baricitinib to restore hair follicle growth in AA patients. Because the IFNγ receptor engages both JAK1 and JAK2, the dual mechanism of action for baricitinib that inhibits both JAK1 and JAK2 is ideally suited for drug selection for patients with AA. The results of the interim analysis of the Phase 2 portion of this study confirm the potential of baricitinib to restore hair growth in patients suffering from AA. By employing a seamless Phase 2/3 study design, rapid progress can be made to fill the current therapeutic void and unmet need.

A washout of systemic and topical treatments for AA is incorporated before randomization to minimize confounding effects of prior treatment. Durations of required washouts take into consideration that several weeks may be necessary before observing regrowth of hair in treated AA patients.

This trial will enroll patients with severe AA (defined as a SALT score of 50% to 95%) and very severe AA (defined as a SALT score of 95% to 100%), as patients with more limited, scattered patches are often treated with topical corticosteroids or intralesional corticosteroid injections (eg, triamcinolone) and topical minoxidil. There is no approved treatment for the management of AA; therefore, this trial will evaluate the efficacy and safety of 2 doses of baricitinib in Phase 3 (baricitinib 2-mg and 4-mg QD) versus placebo.

Eligible patients have at least 6 months of disease with no spontaneous improvement during this period, as spontaneous remissions of AA can be observed during the first months after onset. Hence, minimal chances of spontaneous hair regrowth confounding the ability of baricitinib to facilitate hair follicle regrowth were designed into this study.

The primary endpoint is the proportion of patients reaching a SALT ≤20 (i.e., ≤20% of scalp involvement) and will be evaluated after 36 weeks of treatment. Experience with other drugs has shown that it can take several months for patients with AA under treatment to grow new hair. It is expected that several weeks of treatment will be necessary to allow hair follicles to cycle back

to anagen phase after control of the inflammatory response leading to hair regrowth, and that it could take up to 9 months to observe maximum efficacy.

Alopecia areata is a chronic and relapsing condition, and some patients may require continuous treatment. Patients with SALT ≤20 at Week 52 and who have received the same dose of baricitinib over 52 weeks, will be randomized to either stay on treatment or transition to placebo (in a 3:1 ratio) to observe the consequence of treatment interruption. These patients will be retreated with their initial baricitinib dose if they present with a relapse (loss of >20 points in total SALT score).

Period 4 will provide patients who have completed the Week 104 visit and have not met criteria for permanent discontinuation, the possibility to remain in the trial for up to 96 additional weeks (up to Week 200). This will allow for additional long-term efficacy and safety information to be collected, and provide patients the opportunity to continue study treatment until the anticipated approval of baricitinib in this indication.

5.4.1. Rationale for Use of an Adaptive Operationally Seamless Design

Alopecia areata likely results from loss of the immune privilege of the hair follicle, resulting in its destruction. Several months of treatment may be necessary to observe maximum efficacy with treatment. Therefore, a classical sequential approach of a Phase 2 dose-ranging study followed by two Phase 3 studies would delay the evaluation of the benefit/risk for baricitinib in severe or very severe AA, for which there is currently no approved drug available. Improvements in trial designs that yield more information earlier about the drug may reduce late-stage attrition rates, improve selection of the "right" doses for Phase 3, or terminate development of ineffective or unsafe therapies earlier. Novel clinical trial designs, such as adaptive designs, offer opportunities to address these issues. The Pharmaceutical Research and Manufacturers of America Working Group has expressed that adaptive designs can be utilized beneficially in dose-finding and operationally seamless Phase 2/3 trial designs (Gallo et al. 2006).An adaptive design refers to a clinical trial design that uses accumulating data to decide how to modify aspects of the study as it continues, without undermining the validity and integrity of the trial (Gallo et al. 2006). The goal of adaptation is to learn from the accumulating data and to apply what is learned as quickly as possible. In order to maintain the validity and integrity of the trial, the adaptations must be pre-specified.

Study JAHO, an adaptive, operationally seamless Phase 2/3 trial, is one of two pivotal trials planned for the clinical development of baricitinib in adult patients with severe or very severe AA. A number of factors governed the decision to utilize this design. Based on available preclinical and clinical data, the adaptive design feature is believed to afford more efficient use of patient safety and efficacy data to characterize the baricitinib dose-response relationship and to improve data-driven decisions. The operationally seamless component will allow trial objectives that are normally achieved through separate trials in Phase 2 and 3 to be addressed in a single trial. Specifically, the adaptive dose-finding design should lead to a dose selection decision of the type normally associated with Phase 2. The Phase 3 confirmatory efficacy

objective will be achieved by comparing the safety and efficacy of the selected baricitinib doses to placebo. This operationally seamless design will also provide longer-term safety data earlier in the clinical development program.

If Stage 1 of Study JAHO leads to a decision on baricitinib dose selection, the trial will proceed into the second stage by seamlessly shifting to a fixed-patient allocation scheme, including 2 active arms versus placebo. Patients assigned to the selected baricitinib doses and placebo will remain in the same treatment group. Patients initially assigned to the dose that is discontinued will be transitioned to the highest selected dose. The study schedule will be the same for both stages. Data from patients enrolled before and after adaptation in the baricitinib doses selected and placebo, excluding those enrolled in the Phase 2 portion, will be used for statistical inference in the primary analysis. Therefore, the operationally seamless design will not make inference complicated due to the possibility of type I error inflation which cannot be quantified in a setting of dose modification.

5.5. Justification for Dose

The doses proposed for Study JAHO are baricitinib 1-mg, 2-mg, and 4-mg QD. These doses were chosen primarily based on the recently completed Phase 2 AD study, JAHG, and are additionally supported by PK, safety, and efficacy data for baricitinib in Phase 2 and Phase 3 RA studies and a Phase 2 psoriasis study.

In a Phase 2 study conducted in patients with moderate-to-severe atopic dermatitis (Study I4V-MC-JAHG [JAHG]), both the 2-mg and 4-mg doses showed benefit on the primary and major secondary endpoints, as compared to placebo, and both doses had an acceptable safety profile at Week 16. However, the 4-mg dose appeared to demonstrate a more rapid benefit (at 4 weeks) on more stringent endpoints, compared to the 2-mg dose, particularly in the subgroup of patients with more severe disease at Baseline. The 4-mg dose resulted in statistically significant improvement in the more stringent endpoints at Week 4, and this level of response was maintained through Week 16. A similar trend between the baricitinib 4-mg and 2-mg doses was observed in patients with RA. Although in Study JAHG, the 4-mg dose seemed to perform better than the 2-mg dose on more stringent endpoints, on other endpoints, 2-mg and 4-mg doses showed similar efficacy, compared to placebo. Thus, based on available data, 3 doses will be included in the Phase 2 part of this study, including a 1-mg dose, to cover the range of exposures where clinical responses could be anticipated. The results of the Week 12 interim analysis of the Phase 2 part of Study JAHO showed that for both $SALT_{30}$ (at least 30% improvement from Baseline in SALT score) at Week 12 and $SALT_{50}$ (at least 50% improvement from Baseline in SALT score) at Week 16, the 2-mg and 4-mg baricitinib doses demonstrated numerical superiority over placebo and the 1-mg dose. For $SALT_{50}$ at Week 16, the difference between the 4-mg dose and placebo was statistically significant (p=.036). In the patients with greater disease severity (AA-IGA = 4), the 4-mg dose seemed to provide a more distinct response when compared to the 2-mg dose and the 2-mg dose a more distinct response when compared to the 1-mg dose and placebo, although this observation will need to be confirmed on larger sample sizes. These results were confirmed by the second interim analysis of Phase 2 portion of Study JAHO

conducted after 36 weeks of treatment. There were no SAEs during the period of treatment and no new safety findings.

The pharmacokinetic data from sparse sampling suggest that the systemic exposure in AA patients increases in a dose-dependent manner (approximately dose proportional) across the range from 1-mg to 4-mg. Moreover, the population level exposure of baricitinib for AA patients is similar to that of atopic dermatitis patients and lower than that of rheumatoid arthritis patients.

Altogether, these data support further evaluation of baricitinib 2-mg QD and 4-mg QD regimen in the Phase 3 program of baricitinib in Alopecia Areata, which includes the Phase 3 part of Study JAHO and Study JAIR.

5.5.1. Rationale for Dose Adjustment for Renal Impairment

Baricitinib exposure increases with decreased renal function. Based on PK simulations of baricitinib exposures for the mild and moderate categories of renal function (stratified as eGFR 60 to <90 mL/min/1.73 m^2 and eGFR 30 to <60 mL/min/1.73 m^2, respectively), dose adjustment is not required for patients with eGFR ≥60 mL/min/1.73 m^2. Patients with eGFR <60 mL/min/1.73 m^2 who are randomized to the 4-mg dose will receive a dose of 2-mg QD, which will ensure that exposures do not exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. For patients randomized to the 2-mg dose or 1-mg dose, there will be no dose adjustment based on renal function. The dose adjustment for renal impairment will be managed by IWRS to ensure maintenance of the treatment blind. This study will not enroll patients with screening eGFR <40 mL/min/1.73 m^2. See Section 8.1.1 for eGFR thresholds that trigger interruption of IP.

The procedure of dose adjustment based on renal function (eGFR) during the study is detailed in Section 7.2.2.

6. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria (also known as protocol waiver or exemption) is not permitted.

Study investigator(s) will review patient history and screening test results at Visit 1 and Visit 2 to determine if the patient meets all inclusion criteria and none of the exclusion criteria to qualify for randomization in the study. All screening activities must be completed and reviewed before the patient is randomized.

6.1. Inclusion Criteria

Patients are eligible to be included in the study only if they meet **all** of the following criteria at screening:

Informed Consent

[1] Are at least 18 years and ≤60 years for males (≤70 years of age for females) at the time of informed consent.

Note: Use local requirements to provide consent if the age of adulthood is defined as >18 years. Different upper age limits have been included for male and female patients, based on difference in the prevalence of concomitant androgenetic alopecia.

[2] Are able to read, understand, and give documented (electronic or paper signature) informed consent.

Type of Patient and Disease Characteristics

[3] Have severe or very severe AA, as determined by all of the following:

a. Current AA episode of more than 6 months' duration and hair loss encompassing ≥50% of the scalp, as measured by SALT (AA-IGA of 3 or 4) at Visit 1 AND Visit 2.

b. No spontaneous improvement (i.e., no more than 10 point spontaneous reduction in SALT) over the past 6 months.

c. Current episode of severe or very severe AA of less than 8 years.

Note: patients who have severe or very severe AA for ≥8 years may be enrolled if episodes of regrowth, spontaneous or under treatment, have been observed on the affected areas of the scalp over the past 8 years.

[4] Agree not to use any AA treatments during the study, including, but not limited to:

a. systemic therapies (eg, methotrexate, cyclosporine, corticosteroids, JAK inhibitors, apremilast, dimethyl fumarate derivatives, hydroxychloroquine, mycophenolate-mofetil, IFNγ, azathioprine) and biologics (eg, monoclonal antibodies)

b. intralesional corticosteroid injections

c. topical therapies, including irritants and immunotherapies (eg, diphenylcyclopropenone)

d. Phototherapy, including lasers

e. Platelet-rich plasma injection

f. HMG CoA reductase inhibitors or "Statins" (eg, simvastatin, simvastatin + ezetimibe) for treatment of AA.

g. Cryotherapy

Note: Treatment with bimatoprost ophthalmic solution for eyelashes may be continued if the patient has been on a stable dose for 8 weeks prior to randomization. Treatment with finasteride (or other 5 alpha reductase inhibitors) or oral or topical minoxidil may be continued if the patient has been on a stable dose for 12 months and is anticipated to continue on a stable dose up until Week 36.

[5] Are male or nonpregnant, nonbreastfeeding female patients

a. Male patients will either remain abstinent (if this is their preferred and usual lifestyle) or agree to use 2 forms of birth control (1 must be highly effective, see below) while engaging in sexual intercourse with female partners of childbearing potential and agree to not father a child while enrolled in the study and for at least 4 weeks following the last dose of IP. Men who are in exclusively same-sex relationships (when it is their preferred and usual lifestyle) are not required to use contraception.

b. Female patients of child-bearing potential who are abstinent (if this is complete abstinence, as their preferred and usual lifestyle) or in a same-sex relationship (as part of their preferred and usual lifestyle) must agree to either remain abstinent or stay in a same-sex relationship without sexual relationships with males. Periodic abstinence (eg, calendar, ovulation, symptothermal, post-ovulation methods), declaration of abstinence just for the duration of a trial, and vaginal withdrawal are not acceptable methods of contraception.

   Otherwise, female patients of childbearing potential must agree to use 2 forms of birth control, when engaging in sexual intercourse with male partners while enrolled in the study and for at least 4 weeks following the last dose of IP.

   The following birth control methods are considered acceptable (the patient should choose 2 to be used with their male partners, and 1 must be highly effective):

   - Highly effective birth control methods: oral, injectable, or implanted hormonal contraceptives (combined estrogen/progesterone or progesterone only, associated with inhibition of ovulation); intrauterine device or intrauterine system (eg, progestin-releasing coil); or vasectomized male (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate).

- Effective birth control methods: condom with a spermicidal foam, gel, film, cream, or suppository; occlusive cap (diaphragm or cervical/vault caps) with a spermicidal foam, gel, film, cream, or suppository; or oral hormonal contraceptives that do not inhibit ovulation.

  Note: When local guidelines concerning highly effective or effective methods of birth control differ from the above, the local guidelines must be followed.

c. Females of nonchildbearing potential are not required to use birth control. They are defined as:

- women ≥60 years of age or women who are congenitally sterile, or
- women ≥40 and <60 years of age who have had a cessation of menses for ≥12 months and a follicle-stimulating hormone (FSH) test confirming nonchildbearing potential (≥40 mIU/mL or ≥40 IU/L), or women who are surgically sterile (i.e., have had a hysterectomy or bilateral oophorectomy or tubal ligation).

6.2. Exclusion Criteria

Patients will be excluded from study enrollment if they meet **any** of the following criteria:

Medical Conditions Related to AA

[6] primarily "diffuse" type of AA (characterized by diffuse hair shedding)

[7] are currently experiencing other forms of alopecia, including but not limited to: androgenetic alopecia (male pattern Grade IV or greater using Hamilton-Norwood classification, or female pattern), trichotillomania, telogen effluvium, chemotherapy-induced hair loss or any other concomitant conditions (eg, tinea capitis, psoriasis, lupus erythematosus, or secondary syphilis) that would interfere with evaluations of the effect of study medication on AA.

[8] Patients who, in the opinion of the investigator, are currently experiencing or have a history of unstable concomitant disease that requires frequent hospitalizations and/or frequent use of systemic immunosuppressants that may interfere with participation in the study.

[9] Have been treated with the following therapies:

a. Corticosteroids

i. Topical corticosteroids applied to the scalp or eyebrows within 1 week prior to randomization.

ii. Systemic corticosteroids within 8 weeks prior to randomization.

iii. Intralesional corticosteroid injections for treatment of AA within 8 weeks prior to randomization.

iv. Have had an intraarticular corticosteroid injection within 8 weeks prior to randomization.

Note: Intranasal, ophthalmic, or inhaled steroid use is allowed during screening and throughout the study.

b. JAK inhibitors

i. Topical JAK inhibitor applied to the scalp (eg, tofacitinib, ruxolitinib) within 4 weeks prior to randomization.

ii. Oral JAK inhibitor within 8 weeks prior to randomization.

iii. Previously treated with an oral JAK inhibitor (eg, tofacitinib, ruxolitinib) and had an inadequate response (for example, absence of significant terminal hair growth after at least 12 weeks of treatment).

c. Other topical therapies (eg, anthralin, diphenylcyclopropenone, or other topical immunotherapies) for the treatment of AA within 4 weeks prior to randomization.

d. Monoclonal antibody (eg, ustekinumab, secukinumab, adalimumab, dupilumab) less than 5 half-lives prior to randomization.

e. Probenecid at the time of the randomization (Visit 2) that cannot be discontinued for the duration of the study (probenecid may increase baricitinib exposures).

f. Platelet-rich plasma within 8 weeks prior to randomization.

g. Phototherapy (UV therapy and laser on scalp lesions) within 4 weeks prior to randomization.

h. HMG CoA reductase inhibitors or "Statins" (e.g., simvastatin, simvastatin + ezetimibe) for treatment of AA within 4 weeks prior to randomization.

i. Cryotherapy for treatment of AA within 4 weeks prior to randomization.

j. Finasteride (or other 5 alpha reductase inhibitors) or minoxidil (topical or oral) within 8 weeks prior to randomization, unless the subject has been on a stable dose for at least 12 months AND is anticipated to continue on a stable dose up until Week 36.

k. Immunosuppressants (for example, methotrexate, cyclosporine, dimethyl fumarate derivatives, mycophenolate-mofetil, IFNγ, azathioprine) within 8 weeks of randomization.

l. Apremilast or hydroxychloroquine within 4 weeks prior to randomization.

Medical Conditions in General

[10] Are largely or wholly incapacitated, permitting little or no self-care, such as being bedridden.

[11] Have uncontrolled arterial hypertension characterized by a repeated systolic blood pressure >160 mm Hg or diastolic blood pressure >100 mm Hg in a seated position.

[12] Have had any major surgery within 8 weeks prior to screening or will require major surgery during the study that, in the opinion of the investigator (in consultation with Lilly or its designee), would pose an unacceptable risk to the patient if participating in the trial.

[13] Are immunocompromised and, in the opinion of the investigator, at an unacceptable risk for participating in the study.

[14] Have experienced any of the following within 12 weeks of screening: myocardial infarction, unstable ischemic heart disease, stroke, or New York Heart Association Stage III/IV heart failure.

[15] Have a history of VTE, or are considered at high risk for VTE, as deemed by the investigator, or have 2 or more of the following risk factors for VTE:

a) Aged >65 years

b) BMI >35 kg/m^2

c) Oral contraceptive use and current smoker status

[16] Have a history or presence of cardiovascular, respiratory, hepatic, gastrointestinal, endocrine, hematological, neurological, or neuropsychiatric disorders or any other serious and/or unstable illness that, in the opinion of the investigator, could constitute an unacceptable risk when taking IP or interfere with the interpretation of data.

[17] Have a history of lymphoproliferative disease; or have signs or symptoms suggestive of possible lymphoproliferative disease, including lymphadenopathy or splenomegaly; or have active primary or recurrent malignant disease; or have been in remission from clinically significant malignancy for <5 years.

a. Patients with cervical carcinoma in situ that has been successfully treated with no evidence of recurrence or metastatic disease for at least 3 years may participate in the study.

b. Patients with basal cell or squamous cell skin cancers that have been successfully treated with no evidence of recurrence for at least 3 years may participate in the study.

[18] Have a current or recent and/or serious viral, bacterial, fungal, or parasitic infection, including but not limited to the following:

a. Symptomatic herpes zoster infection within 12 weeks prior to screening.

b. A history of disseminated/complicated herpes zoster (eg, multidermatomal involvement, ophthalmic zoster, central nervous system involvement, or post-herpetic neuralgia).

c. Symptomatic herpes simplex at the time of randomization.

d. Active or chronic viral infection from hepatitis B virus (HBV), hepatitis C virus (HCV), or human immunodeficiency virus (HIV).

e. Household contact with a person with active tuberculosis (TB) and did not receive appropriate and documented prophylaxis for TB.

f. Evidence of active TB or have previously had evidence of active TB and did not receive appropriate and documented treatment.

g. Clinically serious infection, or received intravenous (IV) antibiotics for an infection, within 4 weeks prior to randomization.

h. Any other active or recent infection within 4 weeks of randomization that, in the opinion of the investigator, would pose an unacceptable risk to the patient if participating in the study.

Note: A recent viral upper respiratory tract infection or uncomplicated urinary tract infection should not be considered clinically serious.

[19] A history of eczema herpeticum within 12 months prior to screening.

[20] A history of 2 or more episodes of eczema herpeticum in the past.

[21] Have any serious concomitant illness that is anticipated to require the use of systemic corticosteroids or otherwise interfere with study participation or require active frequent monitoring (eg, atopic dermatitis, unstable chronic asthma).

[22] Have been exposed to a live vaccine within 12 weeks prior to planned randomization or are expected to need/receive a live vaccine during the course of the study (with the exception of herpes zoster vaccination).

Note: Patients eligible for herpes zoster vaccine, who have not received it prior to screening will be encouraged (per local guidelines) to do so prior to randomization; vaccination with the live herpes zoster vaccine can occur during the screening period but must take place >4 weeks prior to randomization and start of IP. Vaccination with the non-live herpes zoster vaccine requires at least 2 injections administered 8 weeks apart. It is recommended that patients who have initiated vaccination with the non-live herpes zoster vaccine receive the second dose at least 4 weeks prior to randomization. For both vaccines, completion of the vaccination may require rescreening the patients to comply with the maximum allowed duration of the screening period. Patients will be excluded if they were exposed to herpes zoster vaccination within 4 weeks of planned randomization.

[23] Have a history of chronic alcohol abuse, IV drug abuse, or other illicit drug abuse within the 2 years prior to screening.

[24] Presence of significant uncontrolled neuropsychiatric disorder, are clinically judged by the investigator to be at risk for suicide, or have a “yes” answer to any of the following:

a. Question 4 (Active Suicidal Ideation with Some Intent to Act, Without Specific Plan) on the “Suicidal Ideation” portion of the Columbia Suicide Severity Rating Scale (C-SSRS); or

b. Question 5 (Active Suicidal Ideation with Specific Plan and Intent) on the "Suicidal Ideation" portion of the C-SSRS; or

c. Any of the suicide-related behaviors (actual attempt, interrupted attempt, aborted attempt, preparatory act or behavior) on the "Suicidal Behavior" portion of the C-SSRS;

and the ideation or behavior occurred within 2 months of Visit 1.

Note: A patient does not necessarily have to be excluded if they have self-injurious behavior that would be classified as nonsuicidal self-injurious behavior. If this situation arises, the patient should be referred to a psychiatrist or appropriately trained professional, as indicated.

[25] Have donated more than a single unit of blood within 4 weeks prior to screening or intend to donate blood during the course of the study.

Diagnostic Assessments

[26] Have screening electrocardiogram (ECG) abnormalities that, in the opinion of the investigator, are clinically significant and indicate an unacceptable risk for the patient's participation in the study.

[27] Have evidence of active TB or latent TB

a. Have evidence of active TB, defined in this study as the following:

i. Documented by a positive PPD test (≥5 mm of induration between approximately 48 and 72 hours after application, regardless of vaccination history), medical history, clinical features, and abnormal chest x-ray at screening.

ii. The QuantiFERON®-TB Gold test or T SPOT.®*TB* test (as available and if compliant with local TB guidelines) may be used instead of the PPD test. Patients are excluded from the study if the test is not negative and there is clinical evidence of active TB.

Exception: Patients with a history of active TB who have documented evidence of appropriate treatment, have no history of re-exposure since their treatment was completed, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria are met. Such patients would not be required to undergo the protocol specific TB testing for PPD, QuantiFERON-TB Gold test, or T SPOT.*TB* test but must have a chest x-ray at screening.

b. Have evidence of untreated/inadequately or inappropriately treated latent TB, defined in this study as the following:

i. Documented to have a positive PPD test (≥5 mm of induration between approximately 48 and 72 hours after application, regardless of vaccination history), no clinical features consistent with active TB, and a chest x-ray with no evidence of active TB at screening; or

ii. PPD test is positive and the patient has no medical history or chest x-ray findings consistent with active TB, the patient may have a QuantiFERON-TB Gold test or T SPOT.*TB* test (as available and if compliant with local TB guidelines). If the test results are not negative, the patient will be considered to have latent TB (for purposes of this study); or

iii. QuantiFERON-TB Gold test or T SPOT.*TB* test (as available and if compliant with local TB guidelines) may be used instead of the PPD test. If the test results are positive, the patient will be considered to have latent TB. If the test is indeterminate, the test may be repeated once within approximately 2 weeks of the initial value. If the repeat test results are again not negative, the patient will be considered to have latent TB (for purposes of this study).

Exception: Patients who have evidence of latent TB may be enrolled if they complete at least 4 weeks of appropriate treatment prior to randomization and agree to complete the remainder of treatment while in the trial.

Exception: Patients with a history of latent TB who have documented evidence of appropriate treatment, have no history of re-exposure since their treatment was completed, and have a screening chest x-ray with no evidence of active TB may be enrolled if other entry criteria are met. Such patients would not be required to undergo the protocol-specific TB testing for PPD, QuantiFERON-TB Gold test, or T SPOT.*TB* test but must have a chest x-ray at screening.

[28] Have a positive test for HBV infection defined as:

a. Positive for hepatitis B surface antigen (HBsAg), or

b. Positive for hepatitis B core antibody (HBcAb) and positive HBV deoxyribonucleic acid (DNA).

Note: Patients who are HBcAb-positive and HBV DNA-negative may be enrolled in the study. Patients who meet these criteria at screening will be identified by the central laboratory and will need to be monitored during the study.

[29] Have HCV infection (positive for anti-hepatitis C antibody with confirmed presence of HCV ribonucleic acid (RNA)

Note: Patients who have documented anti HCV treatment for a past HCV infection AND are HCV RNA negative may be enrolled in the study.

[30] Have evidence of HIV infection and/or positive HIV antibodies.

[31] Have screening laboratory test values, including thyroid-stimulating hormone (TSH), outside the reference range for the population or investigative site that, in the opinion of the investigator, pose an unacceptable risk for the patient's participation in the study.

Note: Patients who are receiving thyroxine as replacement therapy may participate in the study, provided stable therapy has been administered for ≥12 weeks and TSH is within the laboratory's reference range. Patients who are receiving stable thyroxine replacement therapy who have TSH marginally outside the laboratory's normal reference range may participate if the treating physician has documented that the thyroxine replacement therapy is adequate for the patient.

[32] Have any of the following specific abnormalities on screening laboratory tests:

a. AST or ALT ≥2 × the upper limit of normal (ULN)

b. Alkaline phosphatase (ALP) ≥2 × ULN

c. TBL ≥1.5 × ULN

d. Hemoglobin <10.0 g/dL (100.0 g/L)

e. Total white blood cell count <2500 cells/μL(<2.50 × 10^3/μL or <2.50 GI/L)

f. Neutropenia (absolute neutrophil count [ANC] <1200 cells/μL) (<1.20 × 10^3/μL or <1.20 GI/L)

g. Lymphopenia (lymphocyte count <750 cells/μL) (<0.75 × 10^3/μL or <0.75 GI/L)

h. Thrombocytopenia (platelets <100,000/μL) (<100 × 10^3/μL or <100 GI/L)

i. eGFR <40 mL/min/1.73 m^2 (Chronic Kidney Disease Epidemiology Collaboration equation [CKD-EPI] creatinine 2009 equation).

Note: For cases with any of the aforementioned laboratory abnormalities (Exclusion Criteria [31] and [32]), the tests may be repeated during screening, and values resulting from repeat testing may be accepted for enrollment eligibility if they meet the eligibility criterion.

Other Exclusion Criteria

[33] Are unable or unwilling to make themselves available for the duration of the study and/or are unwilling to follow study restrictions/procedures.

[34] Are currently enrolled in any other clinical study involving an investigational product or any other type of medical research judged not to be scientifically or medically compatible with this study.

[35] Have participated within the last 30 days in a clinical study involving an IP. If the previous IP has a long half-life (2 weeks or longer), at least 3 months or 5 half-lives (whichever is longer) should have passed.

[36] Have previously been randomized in this study or any other study investigating baricitinib.

[37] Are investigator site personnel directly affiliated with this study and/or their immediate families. Immediate family is defined as a spouse, parent, child, or sibling, whether biological or legally adopted.

[38] Are Lilly or Incyte employees or their designee.

6.3. Lifestyle Restrictions

Study participants should be instructed not to donate blood or blood products during the study. It is highly recommended that patients remain consistent with hairstyle and hair coloring during the study to facilitate scalp hair assessments. Patients who prefer to shave their scalp must refrain from shaving the scalp within 2 weeks prior to a study visit. For patients who prefer to wear a wig, it is recommended that the wig should not be taped at more than 2 places on the periphery of the scalp to avoid pulling the hair.

Patients with eyebrows and/or eyelashes extensions, and/or tattoos, are allowed to be enrolled in the trial as long as the degree of eyebrows and eyelashes involvement (and regrowth when applicable) can be assessed at every visit. Patients with acrylic nails at screening can be enrolled and will not be requested to remove them to enter the trial; no nails assessment or photographs of the nails will be required for these patients. Patients who do not wear acrylic nails when they enter the trial should refrain from using them before Week 36.

6.4. Screen Failures

Patients who are entered into the study but do not meet the enrollment criteria for participation in this study (screen failure) may be rescreened a maximum of 2 times. The interval between screen failure and rescreenings should be at least 4 weeks. At the time of rescreening, the individual must sign a new informed consent form (ICF), repeat all necessary screening procedures, and will be assigned a new identification number.

If a patient undergoes rescreening, radiographic images acquired as part of initial screening and within 6 months of randomization may be used.

7. Treatments

7.1. Treatments Administered

The Phase 2 portion (included in Stage 1) of this study involves a comparison of placebo, baricitinib 1-mg, baricitinib 2-mg, and baricitinib 4-mg. The Phase 3 portion involves a comparison of placebo, baricitinib 2-mg, and baricitinib 4-mg administered orally QD. Table JAHO.3 shows the treatment regimens and the investigational product (IP) supplied.

The investigator or his/her designee is responsible for the following:

- explaining the correct use of the investigational agent(s) to the patient
- verifying that instructions are followed properly
- maintaining accurate records of IP dispensing and collection
- at the end of the study, returning all unused medication to Lilly or its designee, unless the Sponsor and sites have agreed all unused medication is to be destroyed by the site, as allowed by local law

Table JAHO.3. Treatment Regimen

Treatment Regimen	Investigational Product Supplied	Dose
Baricitinib 4-mg QD[a]	Baricitinib 4-mg tablets Placebo to match 2-mg tablets	2 tablets per day
Baricitinib 2-mg QD	Baricitinib 2-mg tablets Placebo to match 4-mg tablets	2 tablets per day
Placebo QD	Placebo to match 4-mg tablets Placebo to match 2-mg tablets	2 tablets per day

Abbreviation: QD = once daily.

[a] The baricitinib dose for patients randomized to the 4-mg QD treatment group who have renal impairment (defined as eGFR <60 mL/min/1.73 m^2) will be 2-mg QD.

7.1.1. Packaging and Labelling

The Sponsor (or its designee) will provide the following IPs:

- tablets containing 4-mg of baricitinib
- tablets containing 2-mg of baricitinib
- tablets containing 1-mg of baricitinib (only for Phase 2 Population, before Decision Point)
- placebo tablets to match baricitinib 4-mg tablets, 2-mg tablets, and 1-mg tablets (1-mg matching placebo tablets only for Phase 2 Population, before Decision Point)

Each tablet has a distinctive shape and color, 4-mg versus 2-mg versus 1-mg, and each strength tablet has a matching placebo.

During Stage 1:

- The first approximately 100 patients randomized will require 3 double-blind tablets per day and continue this dosing until the Decision Point;
- The remaining approximately 200 patients randomized during Stage 1, will require 2 double-blind tablets per day.

During Stage 2

- After the Decision Point, patients who required 3 double-blind tablets per day during Stage 1 will be switched to 2 double-blind tablets per day;
- Patients randomized during Stage 2 will require 2 double-blind tablets per day.

Clinical trial materials will be labeled according to the country's regulatory requirements.

7.2. Method of Treatment Assignment

Different randomization schemes will be used at Visit 2: two during Stage 1 and one during Stage 2. In Stage 1, the first approximately 100 patients who meet all criteria for enrollment will be randomized in a 1:1:1:1 ratio to receive placebo QD, baricitinib 1-mg QD, baricitinib 2-mg QD, or baricitinib 4-mg QD double-blind treatment at Visit 2 (Week 0). Up to a maximum of approximately 200 additional patients are anticipated to be randomized during Stage 1, prior to the Decision Point, in a 2:2:3 ratio to receive placebo QD, baricitinib 2-mg QD or baricitinib 4-mg QD. At the Decision Point, baricitinib 4-mg and 2-mg were selected to continue into Stage 2. Therefore, after the Decision Point, patients will continue to be randomized in a 2:2:3 ratio to receive placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD. Baseline randomization will be stratified by geographic region (North America, Japan [included in the Phase 2 portion of the study], and North America, Asia, and Rest of World [included in the Phase 3 portion of the study]), and duration of current episode at Baseline (less than 4 years versus at least 4 years) for the whole study. Randomization for the randomized withdrawal will not be stratified. Assignment to treatment groups will be determined by a computer-generated random sequence using an IWRS. The IWRS will be used to assign bottles, each containing double-blind IP tablets, to each patient, starting at Visit 2 (Week 0), and at each visit up to and including Visit 23 (Week 184). Site personnel will confirm that they have located the correct bottles by entering a confirmation number found on the bottle into the IWRS.

7.2.1. Selection and Timing of Doses

The IP should be taken QD without regard to food and, if possible, at approximately the same time every day, usually at the start of the patient's day, to aid patient compliance.

Before Decision Point (Stage 1 patients):

- Patients will take 3 tablets QD for the first approximately100 patients.
- Patients will take 2 tablets QD for the next approximately 200 patients enrolled in Stage 1.

After Decision Point:

- Stage 1 Patients: All patients after the Decision Point will take 2 tablets QD.
- Stage 2 Patients: All patients will take 2 tablets QD.

7.2.2. Dose Adjustment for Renal Impairment

The rationale of dose adjustment for patients with documented renal impairment (defined as screening eGFR ≥40 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2) is detailed in Section 5.5.1.

The dose adjustment for renal impairment will be managed by IWRS to ensure maintenance of the treatment blind. The eGFR value from the screening visit (Visit 1) will be entered into IWRS at Visit 2, and IWRS will assign the treatment doses accordingly.

Patients with documented renal impairment (defined as screening eGFR ≥40 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2), who are randomized to the 4-mg active treatment arm will receive a dose of 2-mg QD by the IWRS. For patients randomized to the 2–mg dose or 1–mg dose, there will be no dose adjustment based on renal function.

No dose adjustment will be made for patients with screening eGFR ≥60 mL/min/1.73 m^2. These patients who are randomized to active treatment will receive their assigned dose, either baricitinib 4-mg, 2-mg, or 1-mg, respectively.

During the study, for patients with documented renal impairment when the subsequent eGFR falls <30 mL/min/1.73 m^2, IP will be withheld until their eGFR becomes ≥40 mL/min/1.73 m^2, whereupon, the IP dosing may resume. For patients with screening eGFR ≥60 mL/min/1.73 m^2, when the subsequent eGFR falls to <40 mL/min/1.73 m^2, IP will be withheld until their eGFR becomes ≥50 mL/min/1.73 m^2, whereupon, the IP dosing may resume (see Section 8.1.1).

7.3. Blinding

This is a double-blind study. To preserve the blinding of the study, a minimum number of Lilly personnel will see the randomization table and treatment assignments before the study is complete. All study assessments will be performed by study personnel who are blinded to the patients' treatment groups. The Decision Point Committee, an internal committee, will be unblinded to the first approximately 100 patients at the Decision Point in Stage 1 but kept blinded to the remaining patients. Processes for unblinding the first approximately 100 patients will be described in the unblinding plan.

Except in clinical circumstances where unblinding is required, the patients, investigators, Lilly study team managing the study, and any other personnel interacting directly with patients or investigative sites will remain blinded to baricitinib and placebo assignment until after completion of the study. It is expected that the need for unblinding a patient's treatment prior to completion of the study will be extremely rare. Every effort should be made to preserve the blind unless there is a compelling reason that knowledge of the specific treatment would alter the medical care of the patient. In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a patient's treatment assignment is warranted for medical management of the event. The patient's safety must always be the first consideration in making such a determination. If a patient's treatment assignment is unblinded, Lilly must be notified

immediately. If the investigator decides that unblinding is warranted, it is the responsibility of the investigator to promptly document the decision and rationale and notify Lilly as soon as possible.

Emergency unblinding for AEs may be performed through the IWRS. This option may be used ONLY if the patient's well-being requires knowledge of the patient's treatment assignment. All unblinding events are recorded and reported by the IWRS. If an investigator, site personnel performing assessments, or patient is unblinded, the patient must be discontinued from the study. In cases where there are ethical reasons to have the patient remain in the study, the investigator must obtain specific approval from a Lilly clinical research physician for the patient to continue in the study.

7.4. Dosage Modification

As described in Section 5.1.6, some patients may be eligible for rescue to baricitinib at Week 36 and transition to baricitinib 4-mg or placebo at Week 52. However, these dose modifications will occur automatically through IWRS and, therefore, will not require dose adjustment by the sites.

7.5. Preparation/Handling/Storage/Accountability

All IP (used and partially used) will be returned to the Sponsor or destroyed at site level with the Sponsor's written approval. In some cases, sites may destroy the material if, during the investigative site selection, the evaluator has verified and documented that the site has appropriate facilities and written procedures to dispose of clinical trial materials.

Follow storage and handling instructions on the IP packaging.

7.6. Treatment Compliance

Patient compliance with study medication will be assessed at each visit during the treatment period (Visit 3 through Visit 24) by counting returned tablets.

A patient will be considered significantly noncompliant if he/she misses more than 20% of the prescribed doses of IP during the study unless the patient's IP is withheld by the investigator for safety reasons. Similarly, a patient will be considered significantly noncompliant if he/she is judged by the investigator to have intentionally or repeatedly taken 20% more than the prescribed amount of medication during the study.

Patients will be counseled by study staff, as appropriate, on the importance of taking the IP as prescribed.

Patients' compliance will be further defined in the statistical analysis plan (SAP).

7.7. Concomitant Therapy

All concomitant medication, whether prescription or over-the-counter, used at Baseline and/or during the course of the study, must be recorded on the Concomitant Medication electronic case report form (eCRF). Patients will be instructed to consult the investigator or other appropriate

study personnel at the site before taking any new medications or supplements during the study. For AA therapies permitted, see Section 7.7.1.

7.7.1. Permitted Medications and Procedures

The following medications are permitted during the study:

- Topical corticosteroids except on the scalp, eyebrows, and eyelids.
- Topical calcineurin inhibitors except on the scalp, eyebrows and eyelids.
- Intranasal, ophthalmic or inhaled steroid use.
- A maximum of 2 intra-articular or soft tissue (bursa, tendon, and/or ligament) corticosteroid injections are allowed up until the 36-week primary endpoint. After 36 weeks, such injections are permitted.
- Non-live vaccinations such as seasonal vaccination, non-live herpes zoster (for subjects who become eligible during the trial), and/or all emergency vaccinations, such as rabies or tetanus vaccinations.
- Bimatoprost ophthalmic solution (if on stable dose for 8 weeks prior to randomization).
- Finasteride (or other 5 alpha reductase inhibitors) or oral or topical minoxidil, if on a stable dose for 12 months prior to randomization.
- HMG CoA reductase inhibitors or "statins" (e.g., simvastatin, simvastatin + ezetimibe) for treatment of hypercholesterolemia and the prevention of cardiovascular disease.

Treatment with concomitant therapies for other medical conditions, such as diabetes and hypertension is permitted during the study.

7.7.2. Prohibited Medications and Procedures

Any investigational or commercial topical, intra-lesional or systemic therapies (except those listed in Section 7.7.1) or phototherapy to treat AA are not allowed during the trial.

In addition, the following medications and procedures are prohibited during the study:

Prohibited Medications and Procedures Requiring Temporary Interruption of Investigational Product

The following therapies will not be allowed during the course of the study and, if taken by or administered to the patient, temporary interruption of IP is required:

- Live vaccines, including Bacillus Calmette-Guérin [BCG] or herpes zoster, (see Exclusion Criterion [22]).
    - For BCG vaccination, IP should be temporarily interrupted for 12 weeks. If BCG vaccine is given prior to Week 36, IP should be permanently discontinued.
    - For live herpes zoster vaccination, IP should be temporarily interrupted for 4 weeks after the injection.

- Probenecid: If a patient is inadvertently started on probenecid, IP should be temporarily interrupted, and can be resumed after patient has discontinued probenecid. If a patient is not able to discontinue probenecid, then IP should be permanently discontinued.
- Systemic corticosteroids may be used for the treatment of an AE (eg, worsening of an existing condition, such as an asthma flare). Investigational product may be restarted if systemic corticosteroids were used for a short duration (<30 days). If used for ≥30 days, Sponsor approval to restart IP is required.
- Phototherapy: Full body UV therapy.

Prohibited Medications Requiring Permanent Discontinuation of Investigational Product

- Corticosteroids (systemic, intralesional, or topical on the scalp, eyebrows, and/or eyelids) for the treatment of AA.
- Topical JAK inhibitors applied to the scalp, eyebrows, and eyelids.
- Other oral JAK inhibitors (eg, tofacitinib and ruxolitinib).
- Any systemic treatment with an immunosuppressive/immunomodulating substance, including, but not limited to, cyclosporine, mycophenolate-mofetil, IFNγ, azathioprine, methotrexate, dimethyl fumarate derivatives, hydroxychloroquine, or biologics (for example, monoclonal antibodies).
- Any other AA treatment which has been inadvertently initiated and cannot be discontinued will lead to permanent discontinuation of IP.

7.8. Treatment after the End of the Study

7.8.1. Continued Access

Period 4 will provide patients who have completed Week 104 visit and have not met criteria for permanent discontinuation, the possibility to remain in the trial for up to 96 additional weeks (up to Week 200) and thus provide patients the opportunity to continue study treatment until the anticipated approval of baricitinib in this indication.

Study termination may occur in a specific country or region when baricitinib is approved for the treatment of AA and becomes reimbursed or commercially available in that country or region, or a negative regulatory action or opinion is received in that country or region.

After the conclusion of the study, continued access to baricitinib will not be provided. Patients will be referred to their local treatment centers for AA therapy, as clinically indicated.

8. Discontinuation Criteria

8.1. Discontinuation from Study Treatment

8.1.1. Permanent Discontinuation from Investigational Product

Investigational product should be permanently discontinued if the patient or the patient's designee requests to discontinue IP.

Discontinuation of the IP for abnormal liver tests should be considered by the investigator when a patient meets 1 of the following conditions after consultation with the Lilly-designated medical monitor:

- ALT or AST >8 × ULN
- ALT or AST >5 × ULN for more than 2 weeks
- ALT or AST >3 × ULN and TBL >2 × ULN or international normalized ratio (INR) >1.5
- ALT or AST >3 × ULN with the appearance of fatigue, nausea, vomiting, right upper-quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)
- ALP >3 × ULN
- ALP >2.5 × ULN and TBL >2 × ULN
- ALP >2.5 × ULN with the appearance of fatigue, nausea, vomiting, right quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)

NOTE: Patients who are discontinued from IP due to a hepatic event or liver test abnormality should have additional hepatic safety data collected via the hepatic safety eCRF.

Investigational product should be permanently discontinued if any of the following are observed:

- white blood cell count <1000 cells/µL (1.00 × 10^3/µL or 1.00 GI/L)
- ANC <500 cells/µL (0.50 × 10^3/µL or 0.50 GI/L)
- lymphocyte count <200 cells/µL (0.20 × 10^3/µL or 0.20 GI/L)
- hemoglobin <6.5 g/dL (<65.0 g/L)

NOTE: Temporary interruption rules (see Section 8.1.2) must be followed, where applicable. For laboratory values that meet permanent discontinuation thresholds, IP should be discontinued. However, if, in the opinion of the investigator, the laboratory abnormality is due to intercurrent illness, such as cholelithiasis or another identified factor, laboratory tests may be repeated. Only when the laboratory value meets resumption thresholds (Table JAHO.4) following the resolution of the intercurrent illness or other identified factor, may the investigator restart IP, after consultation with the Lilly-designated medical monitor.

In addition, patients will be discontinued from IP in the following circumstances:

- pregnancy

- malignancy (except for successfully treated basal or squamous cell skin carcinoma)
- HBV DNA is detected with a value above limit of quantitation (see Section 9.4.8).
- certain prohibited medications are taken per Section 7.7.2 (Prohibited Medications and Procedures)
- development of a VTE

Note: Patients who develop a VTE may have additional follow-up and testing recommended (see Section 9.4.9 and Appendix 7).

If a patient develops multiple risk factors for a VTE during the conduct of the study, as described in Exclusion Criterion [15], the investigator may consider study discontinuation if he/she believes the risk outweighs the benefits of continuing therapy. It is recommended that the investigator consult with Lilly (or its designee) before discontinuing therapy for this reason.

If a patient discontinues IP for any reason, the patient is encouraged to remain in the study through Week 36 (Visit 8) and follow the regular visit schedule to provide the primary efficacy and safety data. Patients discontinuing from the IP prematurely for any reason should complete AE and other follow-up procedures per Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of this protocol.

8.1.2. Temporary Interruption of Investigational Product

In some circumstances, it may be necessary to temporarily interrupt treatment as a result of AEs or abnormal laboratory values that may have an unclear relationship to IP. For example, IP should be temporarily interrupted if the patient experiences a cardiovascular AE considered to be related to study treatment, is graded as moderate (Grade 2 according to Common Terminology Criteria for Adverse Events [CTCAE] Version 3.0), and that does not resolve promptly with supportive care. Except in cases of emergency, it is recommended that the investigator consult with Lilly (or its designee) before temporarily interrupting therapy for reasons other than those defined in Table JAHO.4.

For the abnormal laboratory findings and clinical events (regardless of relatedness) listed in Table JAHO.4, specific guidance is provided for temporarily interrupting treatment and when treatment may be restarted. Retest frequency and timing of follow-up laboratory tests to monitor the abnormal finding are at the discretion of the investigator. Investigational product that was temporarily interrupted because of an AE or abnormal laboratory value not specifically covered in Table JAHO.4 may be restarted at the discretion of the investigator.

Table JAHO.4. Criteria for Temporary Interruption of Investigational Product

Hold Investigational Product if the Following Laboratory Test Results or Clinical Events Occur:	Investigational Product May Be Resumed When:
WBC count <2000 cells/µL (<2.00 × 10^3/µL or <2.00 GI/L)	WBC count ≥2500 cells/µL (≥2.50 × 10^3/µL or ≥2.50 GI/L)
ANC <1000 cells/µL (<1.00 × 10^3/µL or <1.00 GI/L)	ANC ≥1200 cells/µL (≥1.20 × 10^3/µL or ≥1.20 GI/L)
Lymphocyte count <500 cells/µL (<0.50 × 10^3/µL or <0.50 GI/L)	Lymphocyte count ≥750 cells/µL (≥0.75 × 10^3/µL or ≥0.75 GI/L)
Platelet count <75,000/µL (<75 × 10^3/µL or <75 GI/L)	Platelet count ≥100,000/µL (≥100 × 10^3/µL or ≥100 GI/L)
eGFR <40 mL/min/1.73 m^2 (from serum creatinine) for patients with screening eGFR ≥60 mL/min/1.73 m^2	eGFR ≥50 mL/min/1.73 m^2
eGFR <30 mL/min/1.73 m^2 (from serum creatinine) for patients with screening eGFR ≥40 to <60 mL/min/1.73 m^2	eGFR ≥40 mL/min/1.73 m^2
ALT or AST >5 × ULN	ALT and AST return to <2 × ULN, and IP is not considered to be the cause of enzyme elevation
Hemoglobin <8 g/dL (<80.0 g/L)	Hemoglobin ≥10 g/dL (≥100.0 g/L)
Symptomatic herpes zoster	All skin lesions have crusted and are resolving
Infection that, in the opinion of the investigator, merits the IP being interrupted	Resolution of infection
Clinical features of VTE (such as deep vein thrombosis or pulmonary embolism) are present[a]	VTE ruled out

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; eGFR = estimated glomerular filtration rate; IP = investigational product; ULN = upper limit of normal; VTE = venous thromboembolic event; WBC = white blood cell.

[a] Evaluate promptly and institute appropriate treatment. If upon evaluation, VTE is ruled out and no other temporary or permanent discontinuation criteria are met, then IP may be resumed.

Although temporary interruption of IP is not a requirement at times of increased potential risk for VTE (eg, surgery, significant air travel, or other situations involving prolonged immobilization), it is recommended that the investigator follow appropriate VTE prophylaxis guidelines to help manage the VTE risk under these circumstances.

For specific guidance on temporary interruption of IP after use of a prohibited medication, please refer to Section 7.7.2 (Prohibited Medications and Procedures).

Lastly, IP should be temporarily interrupted for suicidal ideation or any suicide-related behaviors, as assessed by the following patient responses on the C-SSRS:

- A "yes" answer to Question 4 (Active Suicidal Ideation with Some Intent to Act, Without Specific Plan); **or**
- A "yes" answer to Question 5 (Active Suicidal Ideation with Specific Plan and Intent) on the "Suicidal Ideation" portion of the C-SSRS; **or**
- A "yes" answer to any of the suicide-related behaviors (actual attempt, interrupted attempt, aborted attempt, preparatory act or behavior) on the "Suicidal Behavior" portion of the C-SSRS.

NOTE: Prior to resumption of IP, it is recommended that a patient be assessed by a psychiatrist or appropriately trained professional to assist in deciding whether the subject should remain on IP and, ultimately, continue participation in the study. A patient does not necessarily have to have IP interrupted if he/she has self-injurious behavior that would be classified as non-suicidal self-injurious behavior.

8.1.3. Discontinuation of Inadvertently Enrolled Patients

If the Sponsor or investigator identify a patient who did not meet enrollment criteria and was inadvertently enrolled, then the patient should be discontinued from study treatment unless there are extenuating circumstances that make it medically necessary for the patient to continue on study treatment. If the investigator and the Sponsor clinical research physician (CRP) agree it is medically appropriate to continue, the investigator must obtain documented approval from the Sponsor CRP to allow the inadvertently enrolled patient to continue in the study with or without treatment with IP. Safety follow-up is as outlined in Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of the protocol.

8.2. Discontinuation from the Study

Patients will be discontinued in the following circumstances:

- enrollment in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- participation in the study needs to be stopped for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and good clinical practice (GCP)
- investigator decision
    - the investigator decides that the patient should be discontinued from the study
    - if the patient, for any reason, requires treatment with another therapeutic agent for AA (not allowed per protocol [Section 7.7.2]) that has been demonstrated to be effective for treatment of the study indication, discontinuation from the study occurs prior to introduction of the new agent
    - if a patient discontinues IP for any reason, the patient is encouraged to remain in the study through Week 36 (Visit 8) and follow the regular visit schedule to provide the primary efficacy and safety data
- subject decision
    - the patient or the patient's designee (eg, parents or legal guardian) requests the patient to be withdrawn from the study

Patients discontinuing from the study prematurely for any reason should complete AE and other safety follow-up per Section 2 (Schedule of Activities), Section 9.2 (Adverse Events), and Section 9.4 (Safety) of the protocol.

8.3. Lost to Follow-Up

A patient will be considered lost to follow-up if he/she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel are expected to make diligent attempts to contact patients who fail to return for a scheduled visit or were, otherwise, unable to be followed up by the site.

9. Study Assessments and Procedures

Section 2 lists the Schedule of Activities, with the study procedures and their timing (including tolerance limits for timing).

Appendix 2 lists the laboratory tests that will be performed for this study.

Unless, otherwise, stated in the subsections below, all samples collected for specified laboratory tests will be destroyed within 60 days of receipt of confirmed test results. Certain samples may be retained for a longer period, if necessary, to comply with applicable laws, regulations, or laboratory certification standards.

9.1. Efficacy Assessments

9.1.1. Primary Efficacy Assessments

9.1.1.1. Severity of Alopecia Tool (SALT) Score

The SALT uses a visual aid showing the division of the scalp hair into 4 areas with the top constituting 40% of total surface, the posterior/back 24%, right side and left side of scalp 18% each. The percentage of hair loss in each area is determined and is multiplied by the percentage of scalp covered by that area. The total sum of the 4 products of each area will give the SALT score, as developed by the National Alopecia Areata Foundation Working Committee (Olsen et al. 2004). Only terminal hair is included in the SALT; vellus hair or any fine downy hair is not taken into account in the SALT scoring process (Olsen et al. 1999, 2004).

9.1.2. Secondary Efficacy Assessments

9.1.2.1. Alopecia Areata Patient-Reported Outcomes

9.1.2.1.1. Patient Reported Outcomes for Scalp Hair Assessment

Lilly has developed a novel patient-reported outcome (PRO) assessment of the patient's current extent of scalp involvement. Like the AA-IGA (see Section 9.1.5.), it is comprised of 5 category response options: 0 = No missing hair (0% of my scalp is missing hair; I have a full head of hair); 1 = A limited area (1% to 20% of my scalp is missing hair); 2 = A moderate area (21% to 49% of my scalp is missing hair); 3 = A large area (50% to 94% of my scalp is missing hair); and 4 = Nearly all or all (95% to 100% of my scalp is missing hair).

9.1.2.1.2. Other Patient Reported Outcomes for Appearance of Eyebrows, Appearance of Eyelashes, Eye Irritation, and Nail Appearance

Lilly has developed 4 other novel PRO assessments measuring 3 important AA signs and 1 important AA symptom: PRO Measure for Eyebrows™, PRO Measure for Eyelashes™, PRO Measure for Eye Irritation™, and PRO Measure for Nail Appearance™. Each of these other PRO assessments uses a 4-point response scale, ranging from 0 = normal appearance/no problem to 3 = severe appearance/severe problem.

9.1.2.2. Other Clinician-Reported Outcomes for Eyebrow Hair Loss Eyelash Hair Loss, and Nail Appearance

Lilly has developed 3 other novel clinician-reported outcome (ClinRO) assessments measuring 3 important AA signs: ClinRO Measure for Eyebrow Hair Loss™, ClinRO Measure for Eyelash Hair Loss™, and ClinRO Measure for Nail Appearance™. Each of these other ClinRO assessments uses a 4-point response scale, ranging from 0 = normal appearance/no hair loss to 3 = severe appearance/severe hair loss.

9.1.3. Health Outcomes and Quality-of-Life Measures

The following patient self-reported questionnaires will be administered via an electronic tablet. In other countries, where necessary, the questionnaires have been translated into the native language of the region and linguistically validated.

9.1.3.1. Skindex-16 Adapted for Alopecia Areata (Stage 2 Only)

Skindex-16 has been used to assess the health-related quality of life in patients with skin diseases. The Skindex-16 items' wordings were adapted for use among adults with AA. It is composed of 16 items grouped under 3 domains: Symptoms (4 items), Emotions (7 items), and Functioning (5 items). The Skindex-16 Adapted for AA will only be used during Stage 2.

9.1.3.2. Medical Outcomes Study 36-Item Short-Form Health Survey Version 2 Acute

The Short-Form Health Survey (SF-36), Version 2, acute measure is a subjective, generic, health-related quality-of-life instrument that is patient-reported and consists of 36 questions covering 8 health domains: physical functioning; bodily pain; role limitations due to physical problems; role limitations due to emotional problems; general health perceptions; mental health; social function; and vitality. Each domain is scored by summing the individual items and transforming the scores into a norm-based T score with higher scores indicating better health status or functioning. In addition, 2 summary scores, the physical component score (PCS) and the mental component score (MCS) will be evaluated based on the 8 SF-36 acute domains (Maruish 2011).

9.1.3.3. EQ-5D-5L

The EQ-5D-5L is a standardized measure of health status 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 respondent's health; and a rating of his/her current health state using a 0 to 100 mm Visual Analog Scale (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 respondent 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. It should be noted that the numerals 1 to 5 have no arithmetic properties and should not be used as an ordinal score. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are labeled "best imaginable health state" and "worst imaginable health state." This information can be used as a quantitative measure of health outcome. The EQ-5D-5L states of health, as defined by the EQ-5D-5L descriptive system, may

be converted into a single summary index by applying a formula that essentially attaches values (also called weights) to each of the levels in each dimension (Herdman et al. 2011; EuroQoL 2015 [WWW]).

9.1.3.4. Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) is a 14-item self-assessment scale that determines the levels of anxiety (7 items) and depression (7 items) that a patient is experiencing over the past week. The HADS utilizes a 4-point Likert response scale (eg, 0 to 3) for each item, and is intended for ages 12 years to 65 years (Zigmond and Snaith 1983; White et al. 1999). Scores for each domain (anxiety and depression) can range from 0 to 21, with higher scores indicating greater anxiety or depression (Zigmond and Snaith 1983; Snaith 2003).

9.1.4. Photography

All sites/patients will obtain photographs of the scalp (4 planes), 1 frontal view of the face and scalp, and photographs of the eyebrows/eyelashes at Baseline (Week 0, Visit 2). Photographs of nails will be obtained at Baseline (Week 0, Visit 2) from patients with nail involvement (ClinRO ≥1) at Baseline, as assessed by the investigator. In addition, photographs of the scalp (4 planes) and 1 photograph (frontal view) of the face and scalp will be obtained at Week 12 (Visit 5), Week 36 (Visit 8, primary endpoint), Week 52 (Visit 11), and at the ETV if patients discontinue prior to Week 52. Photographs of eyebrows/eyelashes and nails will be obtained at Weeks 12, 36, and 52 only from patients with AA involvement (ClinRO ≥1) of these areas at Baseline, as assessed by the investigator. Photographs obtained during the study will provide qualitative visual evidence of clinical response to treatment with baricitinib for all consenting patients. Photographs will not be used for assessments of efficacy, and missed photographs will not be considered a protocol violation.

Camera equipment, necessary ancillary materials, and a study-specific photography manual will be provided to all sites by the Sponsor or designee. The following photographs will be obtained by trained site personnel under similar lighting conditions and magnifications, and per instructions provided during training and as outlined in the photographic procedure manual:

- 4-plane views of the scalp, including each side, back, and top of head.
- 1 frontal view of the face and scalp.
- Photographs of the eyebrows/eyelashes will be taken from patients with eyelash and/or eyebrow loss (ClinRO Measure for Eyebrow and/or Eyelash Hair Loss ≥1).
- A single photograph of each hand will be taken from patients with nail involvement (ClinRO Measure for Nail Appearance ≥1).

Photographs are commonly obtained in research investigations of AA and are effective in demonstrating disease presentation at Baseline and clinical response following treatment with baricitinib.

The photographs may be included in the clinical study report (CSR), a regulatory submission package, scientific publications, or in other public dissemination of clinical data to demonstrate

disease presentation at Baseline and clinical response following treatment with baricitinib. Photographs from consenting patients may also be used in advertising and promotional activities including, but not limited to, communication with healthcare professionals/payers, as well as in patient education, speaker programs, and digital/print media messaging.

Patient anonymity will be protected in photographs included in patient education, speaker programs, print media, or any form available to the public. To protect the patient's anonymity, identifiable characteristics of the skin, such as faces, birthmarks, or tattoos will be redacted from the final photographs and will not appear in any published version of the photograph. No formal analyses of photographs are planned.

9.1.5. Alopecia Areata – Investigator Global Assessment

The AA-IGA is a categorization of overall scalp hair loss based on the patient's SALT score, which is assigned by the investigator by direct inspection of the patient's scalp at each visit (see Section 9.1.1.1). The AA-IGA contains 5 categories: 0 = None (SALT score of 0%); 1 = Limited (SALT score of 1% to 20%); 2 = Moderate (SALT score of 21% to 49%); 3 = Severe (SALT score of 50% to 94%); and 4 = Very Severe (SALT score of 95% to 100%). The AA-IGA will be automatically derived from the SALT score entered into the electronic clinical outcome assessment (eCOA) by the investigator and may be used in exploratory analysis.

9.1.6. Appropriateness of Assessments

All of the clinical and safety assessments in this study are standard, widely used, and generally recognized as reliable, accurate, and relevant. Measurement properties (including reliability and validity) of the newly developed single-item assessments (AA-IGA, ClinROs, and PROs [see Sections 9.1.1 and 9.1.2]) will be evaluated using the Phase 3 trial data.

9.2. Adverse Events

Investigators are responsible for monitoring the safety of patients who have entered this study and for alerting Lilly or its designee to any event that seems unusual, even if this event may be considered an unanticipated benefit to the patient.

The Investigator is responsible for the appropriate medical care of patients during the study.

Investigators must document their review of each laboratory safety report.

The Investigator will record all relevant AE/SAE information in the CRF. The Investigator remains responsible for following, through an appropriate health care option, AEs that are: serious or, otherwise, medically important; considered related to the IP or the study; or that caused the patient to discontinue the IP before completing the study. The patient should be followed until the event resolves, stabilizes with appropriate diagnostic evaluation, or is reasonably explained. The frequency of follow-up evaluations of the AE is left to the discretion of the investigator.

Lack of drug effect is not an AE in clinical studies because the purpose of the clinical study is to establish treatment effect.

After the ICF is signed, study site personnel will record, via eCRF, the occurrence and nature of each patient's preexisting conditions, including clinically significant signs and symptoms of the disease under treatment in the study. In addition, site personnel will record any change in the condition(s) and any new conditions as AEs. Investigators should record their assessment of the potential relatedness of each AE to IP, via eCRF.

The investigator will interpret and document whether an AE has a reasonable possibility of being related to study treatment, to a study device, or to a study procedure, taking into account the disease, concomitant treatment, or pathologies. A "reasonable possibility" means that there is a cause-and-effect relationship between the IP, study device, and/or study procedure and the AE. The investigator answers yes/no when making this assessment.

Planned surgeries and nonsurgical interventions should not be reported as AEs unless the underlying medical condition has worsened during the course of the study.

If a patient's IP is discontinued as a result of an AE, study site personnel must report this to Lilly or its designee via eCRF, clarifying, if possible, the circumstances leading to any dosage modifications or discontinuations of treatment.

9.2.1. Serious Adverse Events

An SAE is any AE from this study that results in one of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threatening experience (i.e., immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- important medical events that may not be immediately life-threatening or result in death or hospitalization, but may jeopardize the patient or may require intervention to prevent one of the other outcomes listed in the definition above. Examples of such medical events include: allergic bronchospasm requiring intensive treatment in an emergency room or at home; blood dyscrasias or convulsions that do not result in inpatient hospitalization; or the development of drug dependency or drug abuse.

All AEs occurring after signing the ICF are recorded in the eCRF and assessed for seriousness criteria. The SAE reporting to the Sponsor begins after the patient has signed the ICF and has received IP. However, if an SAE occurs after signing the ICF, but prior to receiving IP, the SAE should be reported to the Sponsor, as per SAE-reporting requirements and timelines (see Section 9.2) if it is considered reasonably possibly related to study procedure.

Study site personnel must alert Lilly or its designee of any SAE within 24 hours of investigator awareness of the event via a Sponsor-approved method. If alerts are issued via telephone, they are to be immediately followed with official notification on study-specific SAE forms. This 24-hour notification requirement refers to the initial SAE information and all follow-up SAE

information. Patients with a serious hepatic AE should have additional data collected using the eCRF.

Pregnancy (during maternal or paternal exposure to IP) does not meet the definition of an AE. However, to fulfill regulatory requirements, any pregnancy should be reported following the SAE process to collect data on the outcome for both the mother and the fetus.

Investigators are not obligated to actively seek AEs or SAEs in subjects once they have discontinued and/or completed the study (the patient disposition eCRF has been completed). However, if the investigator learns of any SAE, including a death, at any time after a subject has been discharged from the study, and he/she considers the event reasonably possibly related to the study treatment or study participation, the investigator must promptly notify Lilly.

9.2.1.1. Suspected Unexpected Serious Adverse Reactions

Suspected unexpected serious adverse reactions (SUSARs) are serious events that are not listed in the IB and that the investigator identifies as related to IP or procedure. United States 21 CFR 312.32 and European Union Clinical Trial Directive 2001/20/EC and the associated detailed guidances or national regulatory requirements in participating countries require the reporting of SUSARs. Lilly has procedures that will be followed for the identification, recording, and expedited reporting of SUSARs that are consistent with global regulations and the associated detailed guidances.

9.2.2. Adverse Events of Special Interest

Adverse events of special interest will include the following:

- infections (including TB, herpes zoster, or opportunistic infections)
- malignancies
- hepatic events (see Section 9.4.10.1)
- major adverse cardiovascular events (MACE) (see Section 9.4.10)
- thrombotic events (such as deep vein thrombosis [DVT] and pulmonary embolism [PE] and arterial thrombotic event [ATE] (see Section 9.4.9).

Sites will provide details on these AEs, as instructed on the eCRF, and may be asked for additional description by Lilly.

9.2.3. Complaint Handling

Lilly collects product complaints on IPs and drug delivery systems used in clinical studies in order to ensure the safety of study participants, to monitor quality, and to facilitate process and product improvements.

Patients will be instructed to contact the investigator as soon as possible if they have a complaint or problem with the IP so that the situation can be assessed.

9.3. Treatment of Overdose

Refer to the IB.

9.4. Safety

Any clinically significant findings from ECG testing, physical examination, vital signs measurements, or laboratory measurements that result in a diagnosis and that occur after the patient receives the first dose of study treatment should be reported to Lilly or its designee as an AE via eCRF.

9.4.1. Electrocardiograms

A single 12-lead, standard ECG will be obtained locally at Visit 1 and read by a qualified physician (the investigator or qualified designee) at the site to determine whether the patient meets entry criteria.

Electrocardiograms may be obtained at additional times, when deemed clinically necessary.

9.4.2. Vital Signs

For each patient, vital signs should be measured according to the SOA (Section 2).

9.4.3. Physical Examination

For each patient, a complete physical examination (excluding pelvic and rectal examinations) will be performed at Visit 1 (Screening). A symptom-directed physical examination will be performed at other visits, as specified in the SOA (Section 2). A complete physical examination may be repeated, at the investigator's discretion, any time a patient presents with physical complaints.

9.4.4. Laboratory Tests

For each patient, laboratory tests detailed in Appendix 2 should be conducted according to the SOA (Section 2). Home visits to collect blood and urine samples may be allowed upon written approval from the sponsor and if consistent with local regulations. With the exception of laboratory test results that may unblind the study, Lilly or its designee will provide the investigator with the results of laboratory tests analyzed by a central vendor, if a central vendor is used for the clinical trial.

9.4.5. Columbia Suicide Severity Rating Scale

The C-SSRS captures the occurrence, severity, and frequency of suicidal ideation and/or behavior during the assessment period. The scale includes suggested questions to solicit the type of information needed to determine if suicidal ideation and/or behavior occurred. The C-SSRS is administered by an appropriately trained healthcare professional with at least 1 year of patient care/clinical experience. The tool was developed by the National Institute of Mental Health trial group for the purpose of being a counterpart to the Columbia Classification Algorithm of Suicide Assessment categorization of suicidal events. For this study, the scale has been adapted (with

permission from the scale authors) to include only the portion of the scale that captures the occurrence of the 11 preferred ideation and behavior categories.

The nonleading AE collection should occur prior to the collection of the C-SSRS. If a suicide-related event is discovered **during** the C-SSRS, but was not captured during the nonleading AE collection, sites should not change the AE form. If an event is serious or leads to discontinuation, this is an exception where the SAE and/or AE leading to discontinuation should be included on the AE form and the process for reporting SAEs should be followed.

9.4.6. Self-Harm and Follow-Up Supplement Forms

Suicide-related events (behavior and/or ideations) will be assessed and evaluated at every visit with the administration of the C-SSRS and the Self-Harm Supplement Form. The Self-Harm Supplement Form is a single question to enter the number of suicidal behavior events, possible suicide behaviors, or nonsuicidal self-injurious behaviors. If the number of behavioral events is greater than zero, it will lead to the completion of the self-harm follow-up form. The self-harm follow-up form is a series of questions that provides a more detailed description of the behavior cases.

9.4.7. Chest X-Ray and Tuberculosis Testing

A posterior–anterior view chest x-ray will be obtained locally at Screening (Visit 1), unless one has been performed in the past 6 months and the x-ray and/or the report is available. The chest x-ray will be reviewed by the investigator or his/her designee to exclude patients with active TB infection. In addition, patients will be tested at screening (Visit 1) for evidence of active or latent TB, as described in the Exclusion Criteria, Section 6.2.

Investigators should follow local guidelines for monitoring patients for TB if a patient is at high risk for acquiring TB or reactivation of latent TB.

9.4.8. Hepatitis B Virus DNA Monitoring

Hepatitis B virus DNA testing will be performed in enrolled patients who tested positive for HBcAb at screening.

Patients who are HBcAb positive and HBV DNA negative (undetectable) at screening (Visit 1) will require measurement of HBV DNA approximately every 3 months during treatment (as detailed in the SOA [Section 2]) and at the follow-up visit (V801), regardless of their hepatitis B surface antibody (HBsAb) status.

The following actions should be taken in response to HBV DNA test results:

- If a single result is obtained with a value “below limit of quantitation,” the test should be repeated within approximately 2 weeks.
- If the repeat test result is “target not detected,” monitoring will resume according to the study schedule.

- If the patient has 2 or more test results with a value "below limit of quantitation," HBV DNA testing should be performed approximately once per month for the remainder of the study, and referral to a hepatologist is recommended.

  Note: Unscheduled visits may be used to monitor HBV DNA monthly.

- If a result is obtained with a value above the limit of quantitation at any time during the study, the patient will be permanently discontinued from IP (see Section 8.1.1), and should be referred to a hepatology specialist.
  - In selected cases, investigators may temporarily continue IP in accordance with current immunomodulator management in the setting of HBV DNA positivity. This option may be considered in consultation with Lilly (or its designee) and evaluation of individual patient risks and benefits.

9.4.9. Venous Thromboembolism Assessment

If a patient develops the signs and symptoms of a DVT or PE, appropriate local laboratory tests and imaging should be performed, as necessary, for diagnosis of the event. For confirmed cases, additional laboratory testing may be performed, as outlined in Appendix 7. The choice and optimal timing of these tests will be directed by the patient's management and may require ongoing follow-up after study discontinuation. All suspected VTE events will be independently adjudicated by a blinded Clinical Event Committee (see Section 10.3.7.4).

9.4.10. Safety Monitoring

Lilly will periodically review evolving aggregate safety data within the study by appropriate methods.

In the event that safety monitoring uncovers an issue that needs to be addressed by unblinding at the group level, only members of the data monitoring committee (DMC), an advisory group for this study formed to protect the integrity of data, can conduct additional analyses of the safety data (refer to Interim Analyses section [Section 10.3.7]).

The Lilly CRP will monitor safety data throughout the course of the study. Lilly will review SAEs within timeframes mandated by company procedures. The Lilly CRP will, as is appropriate, consult with the functionally independent Global Patient Safety (GPS) therapeutic area physician or clinical scientist and periodically review trends in safety data and laboratory analytes. Any concerning trends in frequency or severity noted by an investigator and/or Lilly (or designee) may require further evaluation.

All deaths and SAE reports will be reviewed in a blinded manner by Lilly during the clinical trial. These reports will be reviewed to ensure completeness and accuracy, but will not be unblinded to Lilly during the clinical trial. If a death or a clinical AE is deemed serious, unexpected, and possibly related to IP, only Lilly GPS will be unblinded for regulatory reporting and safety monitoring purposes. These measures will preserve the integrity of the data collected during this trial and minimize any potential for bias while providing for appropriate safety monitoring.

Investigators will monitor vital signs and carefully review findings that may be associated with cardiovascular events and VTEs. Adverse event reports and vital signs will be collected at each study visit. The cardiovascular monitoring plan includes the following:

- regular monitoring of lipid levels
- potential MACE (cardiovascular death, myocardial infarction, stroke), other cardiovascular events, such as hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock, coronary revascularization (eg, coronary artery bypass graft or percutaneous coronary intervention), venous and arterial thrombotic events, and noncardiovascular deaths will be identified by the investigative site or through medical review and will be sent to a blinded Clinical Event Committee for adjudication at regular intervals.

9.4.10.1. Hepatic Safety Monitoring

If a study patient experiences elevated ALT ≥3 × ULN, ALP ≥2 × ULN, or elevated TBL ≥2 × ULN, liver testing (Appendix 4) should be repeated within 3 to 5 days, including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase to confirm the abnormality and to determine if it is increasing or decreasing. If the abnormality persists or worsens, clinical and laboratory monitoring should be initiated by the investigator and in consultation with the study medical monitor. Monitoring of ALT, AST, TBL, and ALP should continue until levels normalize or return to approximate Baseline levels.

Discontinuation criteria of IPs, either temporary interruption or permanent discontinuation, due to abnormal ALT, AST, TBL, or ALP, are detailed in Section 8.1.

Hepatic Safety Data Collection

Additional safety data should be collected via the hepatic eCRF if 1 or more of the following conditions occur:

- elevation of serum ALT to ≥5 × ULN on 2 or more consecutive blood tests
- elevated serum TBL to ≥ 2 × ULN (except for cases of known Gilbert's syndrome)
- elevation of serum ALP to ≥2 × ULN on 2 or more consecutive blood tests
- patient discontinued from treatment due to a hepatic event or abnormality of liver tests
- hepatic event considered to be a SAE

See Appendix 4 and Appendix 5 for a description of hepatic laboratory values that warrant patient exclusion from the study, temporary or permanent discontinuation of IP, or additional safety collection via the hepatic eCRF.

9.5. Pharmacokinetics

A single venous blood sample will be drawn at the visits and times indicated below. These blood samples will be used to determine the plasma concentrations of baricitinib using a validated liquid chromatography tandem mass spectrometry method. Blood samples that will be used for

other laboratory assessments (eg, chemistry, hematology) may be drawn at approximately the same time as the samples drawn to determine plasma concentrations of baricitinib, with the exception of Week 0, which requires specific postdose sampling times. The inability to collect a PK sample will not be considered a protocol violation. The timing will be as follows:

- At Visit 2 (Week 0), patients will take their IP in the clinic, and PK samples will be drawn 15 minutes and 1 hour postdose. Samples for the other clinical laboratory assessments must be drawn prior to receiving the first dose.
- At Visit 3 (Week 4), patients will be asked to take their IP at home prior to visiting the clinic. The clinic visit should be scheduled so that the blood sample collected during this visit is drawn 2 to 4 hours after the dose is taken at home.
- At Visit 4 (Week 8), patients will be asked to take their IP at home prior to visiting the clinic. The clinic visit should be scheduled so that the blood sample collected during this visit is drawn 4 to 6 hours after the oral dose is taken at home.
- For Visit 5 (Week 12), and Visit 6 (Week 16), patients will be asked to NOT take their IP before visiting the clinic, and a blood sample will be collected at any time predose on the day of the clinic visits. If the patient has taken the oral dose prior to the visit, the sample may be drawn anytime postdose, and the inability to collect a predose sample will not be considered a protocol violation.
- For an ET prior to Visit 6 (Week 16), a sample may be drawn anytime if the last dose of IP was taken within the last 48 hours. After Week 16, if the ET is due to an AE, then a sample should be drawn if the last dose of IP was taken within the last 48 hours of the ET at or before Visit 24 (Week 200).
- In the event of an SAE, up to 2 additional blood samples may be taken at the investigator's discretion. If collected, the PK samples should be collected after the reported event, approximately 6 hours apart and within 24 hours of the patient's last dose.

For visits where PK samples will be collected, the actual date and 24-hour clock time of sample collection, and the date and time of the 2 doses prior to the sample being drawn, should be recorded. For Visits 3 and 4, these 2 doses should be the dose given the morning of the day of sample collection and the dose given the previous day. For Visits 5 and 6, these 2 doses should be the dose given the day before the visit and the dose given the day before that.

This sampling schedule should be followed as closely as possible; however, failure to take PK samples at these specified times will not be considered a protocol violation. If the patient fails to follow the directions for a particular visit, the sample should still be collected at that visit, and the date and 24-hour clock time of sample collection and the date and 24-hour clock time of the 2 doses prior to the sample being drawn should be recorded.

Pharmacokinetic samples from patients receiving baricitinib will be assayed; samples from patients receiving placebo may not be assayed. Pharmacokinetic samples will be kept in storage at a laboratory facility designated by the Sponsor. Pharmacokinetic samples may also be assayed for additional exploratory analyses. Pharmacokinetic results will not be provided to investigative sites until the completion of the study or to the blinded study team until the study has been unblinded.

Bioanalytical samples collected to measure IP concentration will be retained for a maximum of 1 year following last patient visit for the study.

9.6. Pharmacodynamics

Not applicable.

9.7. Pharmacogenomics

9.7.1. Whole Blood Samples for Pharmacogenetic Research

A whole blood sample will be collected for pharmacogenetic analysis, as specified in the SOA (Section 2), where local regulations allow.

Samples will not be used to conduct unspecified disease or population genetic research, either now or in the future. Samples will be used to investigate variable response baricitinib and to investigate genetic variants thought to play a role in AA. Assessment of variable response may include evaluation of AEs or differences in efficacy.

All samples will be coded with the patient number. These samples and any data generated can be linked back to the patient only by the investigator site personnel.

Samples will be retained at a facility selected by Lilly for a maximum of 15 years after the last patient visit for the study, or for a shorter period if local regulations and/or ethical review boards (ERBs)/investigational review boards (IRBs) impose shorter time limits. This retention period enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of baricitinib or after baricitinib become(s) commercially available.

Molecular technologies are expected to improve during the 15-year storage period and, therefore, cannot be specifically named. However, existing approaches include whole genome or exome sequencing, genome-wide association studies, and candidate gene studies. Regardless of technology utilized for genotyping, data generated will be used only for the specific research scope described in this section.

9.8. Biomarkers

Biomarker research is performed to address questions of relevance to drug disposition, target engagement, pharmacodynamics (PD), mechanism of action, variability of patient response (including safety), and clinical outcome. Sample collection is incorporated into clinical studies

to enable examination of these questions through measurement of biomolecules including DNA, RNA, proteins, lipids, and other cellular elements.

Blood samples for biomarker research will be collected at the times specified in the SOA (Section 2), where local regulations allow.

Samples will be used for research on the drug target, disease process, variable response to baricitinib, pathways associated with AA, mechanism of action of baricitinib, and/or research method or in validating diagnostic tools or assay(s) related to AA.

All samples will be coded with the patient number. These samples, and any data generated, can be linked back to the patient only by the investigator site personnel.

Samples will be retained at a facility selected by Lilly for a maximum 15 years after the last patient visit for the study, or for a shorter period if local regulations and ERBs impose shorter time limits. This retention period enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of baricitinib or after baricitinib becomes commercially available.

9.9. Medical Resource Utilization and Health Economics

Health economics will be evaluated in this study utilizing the EQ-5D-5L (see Section 9.1.3). Medical resource utilization parameters will not be evaluated in this study.

10. Statistical Considerations

10.1. Sample Size Determination

Study JAHO will screen approximately 1035 patients in order to enroll approximately 725 patients over Stage 1 and Stage 2. Stage 1 aims to enroll a maximum of approximately 300 patients with the first approximately 100 patients randomized in a 1:1:1:1 ratio to placebo QD, baricitinib 1-mg QD, baricitinib 2-mg QD, or baricitinib 4-mg QD and up to a maximum of an additional 200 patients randomized in a 2:2:3 ratio to placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD. This sample size will yield approximately 100 randomized and treated Phase 2 patients who will have completed Week 12 (Visit 5) or discontinued early, and who will be used for the conduct of the Phase 2, Week 12 interim analysis. The goal of this interim analysis is to select up to 2 doses of baricitinib or to stop for futility based on pre-specified criteria. The sample size of approximately 100 patients is also sufficient to select at least one efficacious dose at least 80% of the time, based on the said pre-specified criteria.

Stage 2 randomization will begin with a 2:2:3 ratio for placebo QD, baricitinib 2-mg QD, or baricitinib 4-mg QD after Decision Point. This study is designed so that approximately 425 patients are randomized in Stage 2. All randomized patients in the Phase 3 portion will be included in the primary efficacy analysis. Hence, approximately 625 patients will be eligible for the primary efficacy analysis. This sample size will provide more than 90% power to test the superiority of baricitinib 4-mg to placebo or the superiority of baricitinib 2-mg to placebo in the primary endpoint (the proportion of patients achieving SALT ≤20 at Week 36) based on a 2-sided Fisher exact test within the graphical testing scheme, at an initial significance level of 0.04 for baricitinib 4-mg and 0.01 for baricitinib 2-mg. The assumptions used for the power calculation are as follows: 30% response rate for baricitinib 4-mg, 20% response rate for baricitinib 2-mg, and 5% response rate for placebo (Kennedy Crispin et al. 2016; Mackay-Wiggan et al. 2016). The initial alpha allocation may be adjusted in the SAP when newer information is obtained on the endpoints that are being tested and will be finalized prior to the primary database lock.

Patients who achieve SALT ≤20 at Week 52 (responders) AND who have remained on the same dose of baricitinib from randomization (Visit 2) to Week 52, will enter the randomized withdrawal, which is meant to evaluate the change in clinical response after treatment withdrawal, and does not account for whether the sample size is sufficient to detect statistical difference between baricitinib and placebo. It is expected that there will be approximately 100 patients eligible for the randomized withdrawal.

10.2. Populations for Analyses

For the purpose of analyses, the following populations are defined:

Population	Description
Phase 2, Week 12 Interim Analysis Set (IAS)	The first approximately 100 randomized and treated patients in Phase 2 portion within Stage 1 who completed Visit 5 (Week 12) assessment or discontinued early. Patients will be analyzed according to the IP to which they were randomized at Baseline (Visit 2).
Phase 2, Week 36 Interim Analysis Set (IAS)	The first approximately 100 randomized and treated patients in Phase 2 portion within Stage 1 who completed Visit 8 (Week 36) assessment or discontinued early. Patients will be analyzed according to the IP to which they were randomized at Baseline (Visit 2).
Full Analysis Set (FAS)	All patients enrolled in Phase 3 portion, and who are randomized to baricitinib 4-mg, baricitinib 2-mg, and placebo treatment arms in both Stages 1 and 2 will be included in the FAS. Patients will be analyzed according to the IP to which they were randomized at Baseline (Visit 2).
Modified Full Analysis Set (mFAS) Population	All patients enrolled in Phase 3 portion, and who are randomized to baricitinib 4-mg, baricitinib 2-mg, and placebo treatment arms in both Stages 1 and 2, and received at least 1 dose of IP, will be included in the mFAS. It excludes patients with female pattern baldness and male patients with diffuse AGA[a] (Grade IV and above) (Norwood 1975) identified at Week 36. Patients will be analyzed according to the IP to which they were randomized at Baseline (Visit 2).
Per-Protocol Set (PPS)	The PPS will include all mFAS patients who are not deemed noncompliant with treatment, who do not have any of the important protocol deviations that exclude patients from the PPS, and whose investigator site does not have significant GCP deviations that require a report to regulatory agencies. The important protocol deviations, including the subset that result in exclusion from the PPS, will be determined while the study team remains blinded, prior to the primary outcome database lock.
Safety Population	The safety population is defined as all randomized patients who receive at least 1 dose of investigational product (IP) and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit. Patients will be analyzed according to the treatment group to which they were assigned.

Abbreviations: AGA = androgenetic alopecia; GCP = good clinical practice; IP = investigational product.

[a] Some male patients with Grade IV AGA and female patients with patterned baldness may only be identified after hair regrowth on the scalp.

Any additional analysis populations will be defined in the SAP.

10.3. Statistical Analyses

10.3.1. General Statistical Considerations

Statistical analysis of this study will be the responsibility of Lilly.

The interim analysis at the Decision Point would be conducted in the Phase 2, Week 12 IAS population. A second interim analysis will be conducted using the Phase 2, Week 36 IAS population. The efficacy analysis of the primary and key secondary endpoints in the Phase 3

portion will be conducted in the FAS population. All other efficacy analyses will be conducted in the FAS population or other populations described in the SAP. Safety analyses will be conducted using the safety population. Additional efficacy or safety interim analyses prior to the final database lock (F-DBL) may occur to support regulatory submissions and will be described in the SAP.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05 unless, otherwise, stated.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol (and the justification for making the change) will be described in the CSR. Additional exploratory analyses of the data will be conducted, as deemed appropriate. Complete details of the planned analyses will be documented in the SAP.

10.3.1.1. Analysis Methods

The primary analysis of discrete efficacy and health outcomes variables will use a logistic regression analysis with geographic region, duration of current episode at Baseline (<4 years vs ≥4 years), baseline value, and treatment group in the model unless, otherwise, stated. The p-value and 95% confidence interval (CI) for the odds ratio from the logistic regression model are used for primary statistical inference. In the case when logistic regression model does not produce statistical results due to sparse data, Fisher exact test will be used. The difference in percentages and 95% CI of the difference in percentages using the Newcombe-Wilson method without continuity correction are used for descriptive purposes unless, otherwise, specified. Missing data will generally be imputed using NRI, as described in Section 10.3.1.2.

The primary analyses for the continuous efficacy and health outcome variables will use ANCOVA with geographic region, duration of current episode at Baseline (<4 years vs ≥4 years), treatment group, and baseline value in the model unless, otherwise, stated. Type III tests for the least-squares (LS) means will be used for statistical comparison between treatment groups. The LS mean difference, standard error, p-value, and 95% CI will also be reported. The method used to handle missing data will be modified last observation carried forward (mLOCF) and is reasonable for this data as very few patients experienced waxing and waning in scalp hair during the course of treatment from the Phase 2 portion and several external studies on alopecia areata. Additional details of the mLOCF method are described in Section 10.3.1.2.

Analysis of continuous efficacy and health outcome variables over time may be made using a restricted maximum likelihood-based mixed model for repeated measures (MMRM) analysis. The model will include geographic region, duration of current episode at Baseline (<4 years vs ≥4 years), treatment group, visit, treatment-by-visit interaction as fixed categorical effects, and baseline value/baseline value-by-visit interaction as fixed continuous effects unless, otherwise, stated. An unstructured covariance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by heterogeneous Toeplitz (TOEPH) will be used. The variance-covariance structure that results in the smallest Akaike

information criterion will be used. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the LS means will be used for the statistical comparisons. The LS mean difference, standard error, p-value, and CIs will also be reported. Contrasts will be set up within the model to test treatment groups at specific time points of interest. Additional details of the MMRM method are described in Section 10.3.1.2.

Time-to-event analysis will be performed and analyzed using log-rank test. A Cox proportional hazards model may be used with treatment and other stratification variables in the model unless, otherwise, stated. Hazard ratio with CIs may be reported. Kaplan–Meier curves will also be produced. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed, and these would be described in detail in the SAP. If the assumption of proportional hazards is not justified, nonproportionality may be modeled by stratification.

The Fisher exact test will be used to perform any between-treatment group comparisons for AEs, discontinuations, and other categorical safety data. Continuous vital signs, body weight, and other continuous safety variables, including laboratory variables, will be analyzed using an ANCOVA with treatment and baseline value in the model. Type III tests will be used for statistical comparison between treatment groups. Shift tables for categorical safety analyses (eg, "high" or "low" laboratory results) will also be produced.

10.3.1.2. Missing Data Imputation

As with any clinical study, patient dropouts and consequently missing data is an expected issue for this study. Patients who discontinue treatment, unless they withdraw consent, will be encouraged to remain in the study and follow the scheduled visits until the primary analysis timepoint at Week 36 to provide the efficacy and safety data needed. While every effort will be made to reduce missing data, the missing data imputation methods described below will be used to provide a conservative approach for assessing efficacy endpoints when patients have missing data.

The following imputation rules will be used for the double-blind treatment period (Period 2). Details for handling missing data in later periods will be provided in the SAP.

- Non-responder Imputation: Analysis of categorical efficacy and health outcomes variables will be assessed using NRI unless, otherwise, stated. Patients will be considered nonresponders for the NRI-based analysis if they do not meet clinical response criteria or if they permanently discontinued study treatment or discontinued from the study at any time prior to the timepoint of interest for any reason. An additional analysis may be performed on all available data up to permanent treatment discontinuation.
- Modified last observation carried forward: For continuous variables, the primary analysis will be ANCOVA with missing data imputed by the mLOCF method, which uses the most recent non-missing post-baseline assessment. The specific modification to the LOCF is data after permanent study treatment discontinuation will not be carried forward.
- Mixed model for repeated measures: For continuous variables, an MMRM analysis with a missing at random assumption for handling missing data may also be conducted. This analysis takes into account both the missingness of data and the correlation of the repeated

measurements. No additional imputation methods will be applied to the MMRM analysis. For patients who permanently discontinue the study or the study treatment, data after discontinuation will be excluded.

Additional analyses, including additional methods for handling missing data, such as placebo multiple imputation (pMI) and tipping point analysis, are specified in Appendix 6 and finalized details (if necessary) may be included in the SAP.

10.3.1.3. Multiple Comparisons/Multiplicity

This study uses an operationally seamless adaptive Phase 2/3 study design. Data collected from the additional approximately 200 patients enrolled in Stage 1 and all patients enrolled in Stage 2 will remain blinded, and the FAS population will be used for the primary efficacy analysis. Pre-specified changes in randomization ratio and allocation to selected doses into Stage 2 will be triggered only through IWRS. Hence, the Type I error rate for the primary efficacy analysis is controlled at a 2-sided alpha level of 0.05.

Multiplicity adjusted analyses will be performed on the primary and key secondary objectives in order to control the overall family-wise Type I error rate at a 2-sided alpha level of 0.05. The graphical multiple testing procedure described in Bretz et al. (2011) will be used. The graphical approach is a closed testing procedure; hence, it strongly controls the familywise error rate across all endpoints (Alosh et al. 2014). Figure JAHO.3 illustrates the graphical testing procedures that will be used; a summary of the procedures is provided in Appendix 6. Details of the specific graphical testing scheme (including testing order, interrelationships, Type I error allocation for the primary and key secondary endpoints, and the associated propagation) will be pre-specified in Appendix 6 and the SAP. These details may be revised in the SAP when newer information is obtained on the endpoints that are being tested. However, the graphical testing scheme will be finalized prior to primary database lock.

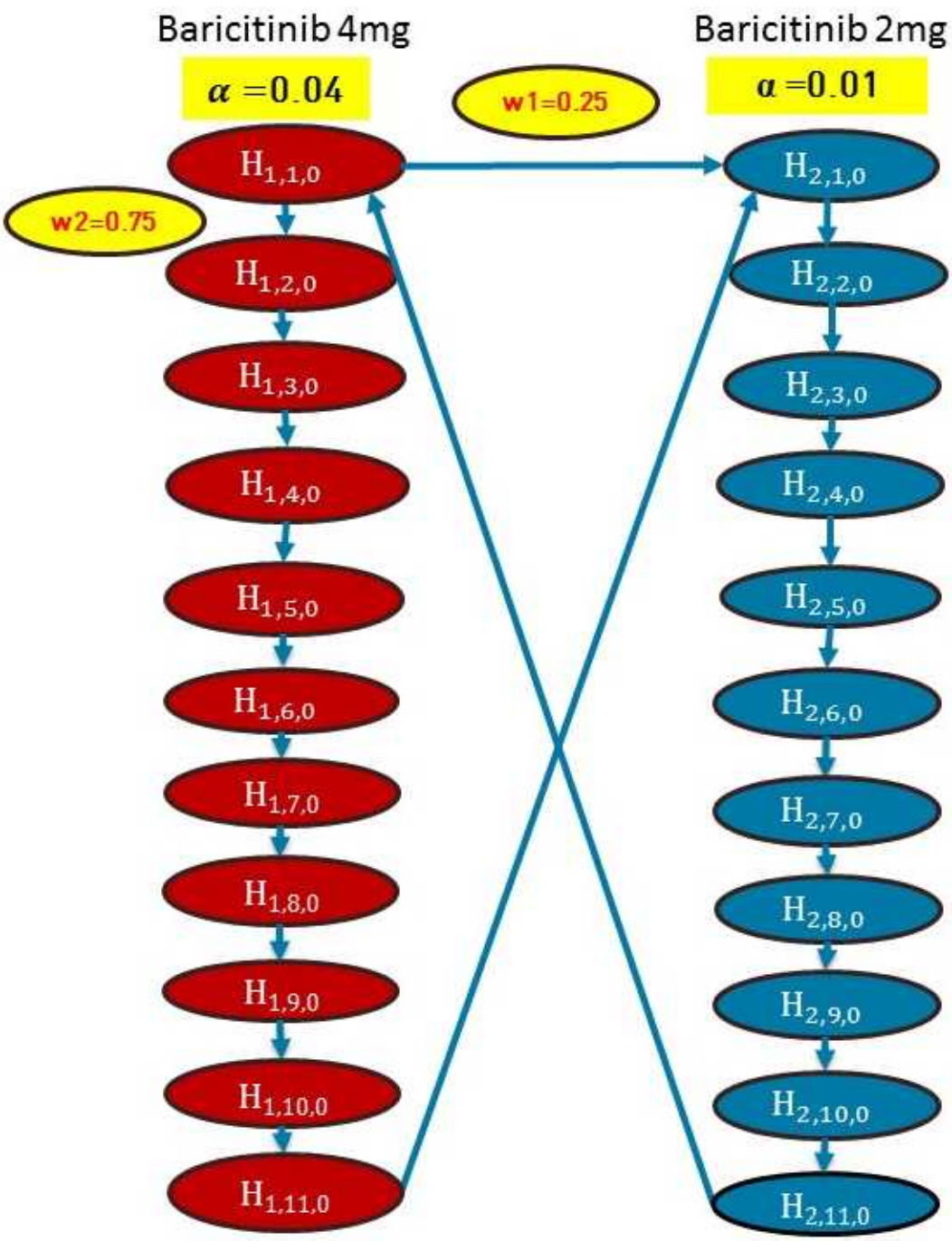


Figure JAHO.3. Graphical testing procedure for I4V-MC-JAHO.

The following is a list of primary and key secondary hypotheses to be tested. The subscript for H denotes the selected 2 doses, the numerical identifier of the endpoint within the dose, and the type of hypothesis (0 for null, 1 for alternative), respectively.

Primary Null Hypothesis:

- $H_{1,1,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving SALT ≤20 at Week 36 is less than or equal to placebo.
- $H_{2,1,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving SALT ≤20 at Week 36 is less than or equal to placebo.

Key Secondaries Null Hypotheses:

- $H_{1,2,0}$: Proportion of patients in the baricitinib 4-mg dose group with a PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement at Week 36 is less than or equal to placebo (among patients with a score of ≥3 at baseline).
- $H_{2,2,0}$: Proportion of patients in the baricitinib 2-mg dose group with a PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement at Week 36 is less than or equal to placebo (among patients with a score of ≥3 at baseline).
- $H_{1,3,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving SALT ≤20 at Week 24 is less than or equal to placebo.
- $H_{2,3,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving SALT ≤20 at Week 24 is less than or equal to placebo.
- $H_{1,4,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving SALT ≤20 at Week 16 is less than or equal to placebo.
- $H_{2,4,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving SALT ≤20 at Week 16 is less than or equal to placebo.
- $H_{1,5,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving an absolute SALT score ≤10 at Week 36 is less than or equal to placebo.
- $H_{2,5,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving an absolute SALT score ≤10 at Week 36 is less than or equal to placebo.
- $H_{1,6,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving an absolute SALT score ≤10 at Week 24 is less than or equal to placebo.
- $H_{2,6,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving an absolute SALT score ≤10 at Week 24 is less than or equal to placebo.
- $H_{1,7,0}$: Proportion of patients in the baricitinib 4-mg dose group with a $SALT_{90}$ at Week 36 is less than or equal to placebo.
- $H_{2,7,0}$: Proportion of patients in the baricitinib 2-mg dose group with a $SALT_{90}$ at Week 36 is less than or equal to placebo.
- $H_{1,8,0}$: Proportion of patients in the baricitinib 4-mg dose group with a $SALT_{50}$ at Week 12 is less than or equal to placebo.
- $H_{2,8,0}$: Proportion of patients in the baricitinib 2-mg dose group with a $SALT_{50}$ at Week 12 is less than or equal to placebo.
- $H_{1,9,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 is less than or equal to placebo (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline)

- $H_{2,9,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 is less than or equal to placebo (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline)
- $H_{1,10,0}$: Proportion of patients in the baricitinib 4-mg dose group achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 is less than or equal to placebo (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline)
- $H_{2,10,0}$: Proportion of patients in the baricitinib 2-mg dose group achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 is less than or equal to placebo (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline)
- $H_{1,11,0}$: Percent change from Baseline in SALT score of patients in the baricitinib 4-mg dose group at Week 36 is greater than or equal to placebo.
- $H_{2,11,0}$: Percent change from Baseline in SALT score of patients in the baricitinib 2-mg dose group at Week 36 is greater than or equal to placebo.

10.3.2. Treatment Group Comparability

10.3.2.1. Patient Disposition

A detailed description of patient disposition by treatment will be summarized with reasons for discontinuation. Frequency counts and percentages will be presented for each treatment group. All patients who discontinue from the study or the study treatment will be identified, along with their reason for discontinuation.

10.3.2.2. Patient Characteristics

Demographic and Baseline characteristics will be summarized descriptively by treatment group. Descriptive statistics including number of patients, mean, standard deviation, median, minimum, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures. No formal statistical comparisons will be made among treatment groups unless, otherwise, stated. A complete list of patient characteristics and Baseline clinical measures will be provided in the SAP.

10.3.2.3. Concomitant Therapy

Concomitant medications will be descriptively summarized by treatment group in terms of frequencies and percentages using the IAS population for the Phase 2 portion and FAS population for the Phase 3 portion. The medications will be coded accordingly.

10.3.2.4. Treatment Compliance

Treatment compliance with the randomly assigned study medication will be evaluated at every clinic visit through the counts of returned IP tablets. A patient will be considered significantly noncompliant if he/she misses more than 20% of the prescribed doses during the study, unless the patient's IP is withheld by the investigator for safety reasons (i.e., compliance <80%). Similarly, a patient will be considered significantly noncompliant if he/she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of

medication (i.e., compliance ≥120%). Patients who are significantly noncompliant will be excluded from the per-protocol set (PPS) (see Section 10.2).

10.3.3. Efficacy Analyses

10.3.3.1. Primary Analyses

The primary efficacy measure is the binary outcome of response defined as SALT ≤20 (see Section 9.1.1) at Week 36. The primary analysis testing the treatment difference between baricitinib 4-mg and placebo or baricitinib 2-mg and placebo will be conducted using a logistic regression model, as described in Section 10.3.1.1. In the case when logistic regression model does not produce statistical results due to sparse data, Fisher exact test will be used.
As mentioned in Section 10.3.1.3, Type I error is controlled at 0.05 level because no data that is used in the interim analyses will be included in the primary efficacy analysis. The NRI method will be used for missing data, as described in Section 10.3.1.2, as it is a conservative imputation approach for handling missing data for patients who permanently discontinued study treatment or discontinued from the study before primary endpoint due to reasons other than lack of efficacy (e.g., lost to follow-up).

Additional analysis of the primary efficacy outcome may be conducted and include analyzing all available data up to the permanent treatment discontinuation.

10.3.3.2. Secondary Analyses

Treatment comparisons in the proportion of patients achieving a binary response will be analyzed using the logistic regression model defined in Section 10.3.1.1. In the case when logistic regression model does not produce statistical results due to sparse data, Fisher exact test will be used. For binary responses, missing data will be imputed using the NRI method described in Section 10.3.1.2 unless, otherwise, stated. Treatment comparisons in the continuous measures will be analyzed using the ANCOVA method defined in Section 10.3.1.1 unless, otherwise, specified. Time-to-event data will be analyzed using log-rank test along with Kaplan-Meier curves.

Additional analysis of the secondary efficacy outcome may be conducted and include analyzing all available data up to the permanent treatment discontinuation.

10.3.3.3. Exploratory Analyses

Analyses will be conducted for the other exploratory objectives defined in Section 4. Specific details of analyses will be specified in the SAP.

10.3.4. Safety Analyses

All safety data will be descriptively summarized by treatment groups and analyzed using the safety population unless, otherwise, stated.

Treatment-emergent adverse events are defined as AEs that first occurred or worsened in severity after the first dose of study treatment. The number of TEAEs, as well as the number and percentage of patients who experienced at least 1 TEAE, will be summarized using the Medical Dictionary for Regulatory Activities (MedDRA) for each system organ class (or a body system)

and each preferred term by treatment group. Serious adverse events and AEs that lead to discontinuation of IP will also be summarized by treatment group. The Fisher exact test will be used to perform comparisons between each baricitinib dose and the placebo group.

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside of normal reference ranges will be flagged in data listings. Quantitative clinical hematology, chemistry, and urinalysis variables obtained at the Baseline to postbaseline visits will be summarized as changes from Baseline by treatment group and analyzed using ANCOVA with treatment and baseline value in the model. Categorical variables, including the incidence of abnormal values and incidence of AEs of special interest, will be summarized by frequency and percentage of patients in corresponding categories. Shift tables will be presented for selected measures.

Observed values and changes from Baseline for vital signs and physical characteristics will be descriptively summarized by treatment group and timepoint. Change from Baseline in vital signs and body weight will be analyzed using ANCOVA with treatment and baseline value in the model.

Summary tables or listings for the C-SSRS and the Self-Harm and Follow-Up Supplement Forms will be produced, as needed.

The incidence and average duration of IP interruptions will be summarized and compared descriptively among treatment groups. Various techniques may be used to estimate the effects of IP interruptions on safety measures. Further analyses may be performed and will be planned in the SAP.

Full details of the safety analyses and any derivations of AEs of special interest will be documented in the program safety analysis plan and study SAP.

10.3.5. Pharmacokinetic Analyses

All plasma baricitinib concentration-time data will be pooled and evaluated using population PK methods. A covariate screen of patient and study-specific factors will be included in the analyses, based on factors investigated in previous and (if any) ongoing PK analyses, and on their relevance to the target population. Exploratory and/or model-based analyses examining the relationships between baricitinib exposure and efficacy and response endpoints will be conducted. Other analyses of efficacy and safety outcome measures may also be assessed, as scientifically appropriate and warranted by available data. Details about the analyses to be conducted will be contained in the PK/PD analysis plan.

10.3.6. Other Analyses

10.3.6.1. Subgroup Analyses

To assess whether the treatment effect is similar across subgroups for the primary efficacy outcome, a logistic model will be used and will include treatment, baseline value, stratification variables, the subgroup variable (eg, sex), and the subgroup-by-treatment interaction. If the

interaction is statistically significant at alpha=0.10, the nature of the interaction will be explored, that is, within each subgroup, the treatment effect will be estimated.

Subgroups to be evaluated will include: demographic characteristics (such as region or country, gender, age, race), renal function, and Baseline disease characteristics. Additional subgroups, as well as detailed descriptions of the subgroup variables, will be defined in the SAP.

Further definitions for the levels of the subgroup variables, the analysis methodology, and any additional subgroup analyses will be defined in the SAP. Because this study is not adequately powered for subgroup analyses, subgroup analyses will generally be treated as exploratory.

10.3.7. Interim Analyses

10.3.7.1. Decision Point Committee

A Decision Point will occur when first approximately 100 randomized and treated patients have completed their assessments at Week 12 (Visit 5) or discontinued early. At the Decision Point, a Decision Point Committee will review efficacy and safety data and provide a recommendation, based on pre-specified criteria, for which up to 2 doses will advance into Stage 2 or to recommend not to proceed to Stage 2 for futility. If the study continues into Stage 2, the study sites will be informed of the selected baricitinib doses. The study team will also be informed of the selected doses to trigger the conduct of an additional Phase 3 trial, Study JAIR.

Because data collection from patients in Phase 2 portion will still be ongoing, even after the Decision Point, information that may unblind the patients during and after the analyses will not be reported to study sites or blinded study team until the study is complete. Additionally, all randomized patients in Stage 1, other than those enrolled in Phase 2 portion, will remain blinded. These patients will be combined with patients in Stage 2 for primary efficacy analysis.

Unblinding and operation of the Decision Point Committee details will be specified in a separate unblinding plan document.

10.3.7.2. Data Monitoring Committee

A DMC will oversee the conduct of this trial. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential and identified risks for this molecule and for this class of compounds. Data Monitoring Committee membership will include, at a minimum, specialists with expertise in dermatology, statistics, and other appropriate specialties.

The DMC will be authorized to review unblinded results of analyses by treatment group prior to final database lock, including study discontinuation data, AEs/SAEs, clinical laboratory data, vital sign data, etc. The DMC may recommend: continuation of the study, as designed; temporary suspension of enrollment; or the discontinuation of a particular dose regimen or the entire study. The DMC may request to review efficacy data to investigate the benefit/risk relationship in the context of safety observations for ongoing patients in the study, however, the study will not be stopped for positive efficacy results. Details of the DMC will be documented in a DMC charter and DMC analysis plan.

The DMC is authorized to evaluate unblinded interim efficacy and safety analyses during both stages of the study, whereas the Decision Point Committee will only review efficacy and safety data of the Phase 2 portion of Stage 1 and will be blinded to all data that will be used in Phase 3. Study sites will receive information about interim results if they need to know for the dose change or safety of their patients.

Unblinding details will be specified in a separate unblinding plan document.

10.3.7.3. Other Interim Analyses

10.3.7.3.1. Phase 2 Week 36 Interim Analysis

When all patients enrolled in the Phase 2 portion will have completed their assessments at Week 36 (Visit 8) or discontinue early, a second interim analysis of the Phase 2 portion will take place.

Unblinding details will be specified in a separate unblinding plan document

This Phase 2 Week 36 interim analysis may trigger an interim analysis of the Phase 3 portion, before the primary outcome database lock (PO-DBL), for evaluation of futility.

10.3.7.3.2. Week 36 Primary Outcome Analysis and Other Regulatory Submission Activities

- After all randomized patients in the Phase 3 portion complete the primary efficacy assessment at Week 36 (Visit 8) or discontinue early, the database will be locked and data will be unblinded to a limited number of pre-identified individuals to initiate work for submission. Although it is called an interim analysis with respect to the entire Phase 3 population, the primary outcome database lock (PO-DBL) interim analysis is the only and final analysis for the primary endpoint. Therefore, no alpha adjustment for this interim analysis is planned. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded.
- Besides the Decision Point Committee and DMC members, a limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the PO-DBL, to initiate the final population PK/PD model development processes or to initiate work for regulatory submission.
- Another interim analysis will occur for the 3-4-month safety update database lock.
- Additional efficacy or safety interim analyses prior to the final database lock (F-DBL) may occur to support regulatory submissions and scientific disclosures.

If an unplanned interim analysis is deemed necessary, the appropriate Lilly medical director or designee will be consulted to determine whether it is necessary to amend the protocol.

10.3.7.4. Adjudication Committee

A blinded Clinical Event Committee will adjudicate potential MACE (cardiovascular death, myocardial infarction, stroke), other cardiovascular events (such as hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock, coronary revascularization [eg, coronary artery bypass graft or percutaneous coronary intervention]), venous and arterial thrombotic events, and noncardiovascular deaths.

Details of membership, operations, recommendations from the Committee, and the communication plan will be documented in the Charter.

11. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713.

Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J.* 2011;53(6):894-913.

Clark JD, Flanagan ME, Telliez JB. Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. *J Med Chem*. 2014;57(12):5023-5038.

Colón EA, Popkin MK, Callies AL, Dessert NJ, Hordinsky MK. Lifetime prevalence of psychiatric disorders in patients with alopecia areata. *Compr Psychiatry.* 1991;32(3):245-251.

Craiglow BG, Liu LY, King BA. Tofacitinib for the treatment of alopecia areata and variants in adolescents. *J Am Acad Dermatol.* 2017;76(1):29-32.

EuroQol Group. EQ-5D-5L User Guide. Version 2.1. Available at: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf. April 2015. Accessed March 2017.

Fridman JS, Scherle PA, Collins R, Burn TC, Li Y, Li J, Covington MB, Thomas B, Collier P, Favata MF, Wen X, Shi J, McGee R, Haley PJ, Shepard S, Rodgers JD, Yeleswaram S, Hollis G, Newton RC, Metcalf B, Friedman SM, Vaddi K. Selective inhibition of JAK1 and JAK2 is efficacious in rodent models of arthritis: preclinical characterization of INCB028050. *J Immunol*. 2010;184(9):5298-5307.

Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, Pinheiro J; PhRMA Working Group. Adaptive designs in clinical drug development – an executive summary of the PhRMA Working Group. *J Biopharm Stat.* 2006;16(3):275-283.

Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736.

Hordinsky M, Donati A. Alopecia areata: an evidence-based treatment update. *Am J Clin Dermatol.* 2014;15(3):231-246.

Hordinsky M, Junqueira AL. Alopecia areata update. *Semin Cutan Med Surg*. 2015;34(2):72-75.

Hunt N, McHale S. The psychological impact of alopecia. *BMJ.* 2005;331(7522):951-953.

Islam N, Leung PS, Huntley AC, Gershwin ME. The autoimmune basis of alopecia areata: a comprehensive review. *Autoimmun Rev*. 2015;14(2):81-89.

Jabbari A, Dai Z, Xing L, Cerise JE, Ramot Y, Berkun Y, Sanchez GA, Goldbach-Mansky R, Christiano AM, Clynes R, Zlotogorski A. Reversal of alopecia areata following treatment with the JAK1/2 inhibitor baricitinib. *EBioMedicine.* 2015;2(4):351-355.

Jabbari A, Sansaricq F, Cerise J, Chen JC, Bitterman A, Ulerio G, Borbon J, Clynes R, Christiano AM, Mackay-Wiggan J. An open-label pilot study to evaluate the efficacy of tofacitinib in moderate to severe patch-type alopecia areata, totalis, and universalis. *J Invest Dermatol*. 2018 Feb 13. pii: S0022-202X(18)30117-9. [Epub ahead of print]

Janković S, Perić J, Maksimović N, Ćirković A, Marinković J, Janković J, Reljić V, Medenica L. Quality of life in patients with alopecia areata: a hospital-based cross-sectional study. *J Eur Acad Dermatol Venereol.* 2016;30(5):840-846.

Karimkhani C, Boyers LN, Naghavi M, Coffeng LE, Lott JP, Wulf S, Hay R, Williams HC, Weinstock MA, Norton SA, Armstrong AW, Dunnick CA, Norris DA, Dellavalle RP. The global burden of disease associated with alopecia areata. *Br J Dermatol.* 2015;172(5):1424-1426.

Kennedy Crispin M, Ko JM, Craiglow BG, Li S, Shankar G, Urban JR, Chen JC, Cerise JE, Jabbari A, Winge MC, Marinkovich MP, Christiano AM, Oro AE, King BA. Safety and efficacy of the JAK inhibitor tofacitinib citrate in patients with alopecia areata. *JCI Insight.* 2016;1(15):e89776.

Korta DZ, Christiano AM, Bergfeld W, Duvic M, Ellison A, Fu J, Harris JE, Hordinsky MK, King B, Kranz D, Mackay-Wiggan J, McMichael A, Norris DA, Price V, Shapiro J, Atanaskova Mesinkovska N. Alopecia areata is a medical disease. *J Am Acad Dermatol.* 2018;78(4):832-834.

Krakowski AC, Eichenfield LF, Dohil MA. Management of atopic dermatitis in the pediatric population. *Pediatrics*. 2008;122(4):812-824.

Liu LY, Craiglow BG, Dai F, King BA. Tofacitinib for the treatment of severe alopecia areata and variants: a study of 90 patients. *J Am Acad Dermatol.* 2017;76(1):22-28.

Liu LY, King BA. Tofacitinib for the treatment of severe alopecia areata in adults and adolescents. *J Investig Dermatol Symp Proc*. 2018;19(1):S18-S20.

Liu LY, King BA, Craiglow BG. Health-related quality of life (HRQoL) among patients with alopecia areata (AA): a systematic review. *J Am Acad Dermatol*. 2016;75(4):806-812.

Mackay-Wiggan J, Jabbari A, Nguyen N, Cerise JE, Clark C, Ulerio G, Furniss M, Vaughan R, Christiano AM, Clynes R. Oral ruxolitinib induces hair regrowth in patients with moderate-to-severe alopecia areata. *JCI Insight.* 2016;1(15):e89790.

Maruish ME. User's manual for the SF-36v2 Health Survey (3rd ed.). Lincoln, RI: QualityMetric Incorporated. 2011.

Norwood OT. Male pattern baldness: classification and incidence. *South Med J.* 1975;68(11):1359–1365.

Olsen EA. Investigative guidelines for alopecia areata. *Dermatol Ther*. 2011;24(3):311-319.

Olsen EA, Bettencourt MS, Coté NL. The presence of loose anagen hairs obtained by hair pull in the normal population. *J Investig Dermatol Symp Proc*. 1999;4(3):258-260.

Olsen EA, Hordinsky MK, Price VH, Roberts JL, Shapiro J, Canfield D, Duvic M, King LE Jr, McMichael AJ, Randall VA, Turner ML, Sperling L, Whiting DA, Norris D; National Alopecia Areata Foundation. Alopecia areata investigational assessment guidelines – Part II. National Alopecia Areata Foundation. *J Am Acad Dermatol.* 2004;51(3):440-447.

Paus R, Nickoloff BJ, Ito T. A 'hairy' privilege. *Trends Immunol*. 2005;26(1):32-40.

Rencz F, Gulácsi L, Péntek M, Wikonkál N, Baji P, Brodszky V. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. *Br J Dermatol.* 2016;175(3):561-571.

Snaith RP. The Hospital Anxiety and Depression Scale. *Health Qual Life Outcomes*. 2003;1:29.

Strazzulla LC, Wang EH, Avila L, Lo Sicco K, Brinster N, Christiano AM, Shapiro J. Alopecia areata: disease characteristics, clinical evaluation, and new perspectives on pathogenesis. *J Am Acad Dermatol.* 2018;78(1):1-12.

Suárez-Fariñas M, Ungar B, Noda S, Shroff A, Mansouri Y, Fuentes-Duculan J, Czernik A, Zheng X, Estrada YD, Xu H, Peng X, Shemer A, Krueger JG, Lebwohl MG, Guttman-Yassky E. Alopecia areata profiling shows TH1, TH2, and IL-23 cytokine activation without parallel TH17/TH22 skewing. *J Allergy Clin Immunol*. 2015;136(5):1277-1287.

Tan E, Tay YK, Goh CL, Giam Y. The pattern and profile of alopecia areata in Singapore – a study of 219 Asians. *Int J Dermatol.* 2002;41(11):748-753.

Villasante Fricke AC, Miteva M. Epidemiology and burden of alopecia areata: a systematic review. *Clin Cosmet Investig Dermatol.* 2015;8:397-403.

Wasserman D, Guzman-Sanchez DA, Scott K, McMichael A. Alopecia areata. *Int J Dermatol.* 2007;46(2):121-131.

White D, Leach C, Sims R, Atkinson M, Cottrell D. Validation of the Hospital Anxiety and Depression Scale for use with adolescents. *Br J Psychiatry*. 1999;175:452-454.

Xing L, Dai Z, Jabbari A, Cerise JE, Higgins CA, Gong W, de Jong A, Harel S, DeStefano GM, Rothman L, Singh P, Petukhova L, Mackay-Wiggan J, Christiano AM, Clynes R. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. *Nat Med.* 2014;20(9):1043-1049.

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361-370.

12. Appendices

Appendix 1. Abbreviations and Definitions

Term	Definition
AA	alopecia areata
AA-IGA	Alopecia Areata Investigator Global Assessment
AD	atopic dermatitis
AE	adverse event: any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product that does not necessarily have a causal relationship with this treatment. An adverse event can, therefore, be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
APC	activated protein C
APTT	activated partial thromboplastin time
AST	aspartate aminotransferase
AT	alopecia areata totalis
ATE	arterial thrombotic event
AU	alopecia areata universalis
BCG	Bacillus Calmette-Guérin vaccine
blinding/masking	A single-blind study is one in which the investigator and/or his staff are aware of the treatment but the patient is not, or vice versa, or when the Sponsor is aware of the treatment but the investigator and/his staff and the patient are not. A double-blind study is one in which neither the patient nor any of the investigator or Sponsor staff who are involved in the treatment or clinical evaluation of the subjects are aware of the treatment received.
BMI	body mass index
BUN	blood urea nitrogen
CI	confidence interval

CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
ClinRO	clinician-reported outcome
CPK	creatinine phosphokinase
CRF	case report form
CRP	clinical research physician
CSR	clinical study report
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DVT	deep vein thrombosis
ECG	electrocardiogram
eCOA	electronic clinical outcome assessment
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
ePRO/PRO	electronic patient-reported outcome/patient-reported outcome
EQ-5D-5L	European Quality of Life – 5 Dimensions – 5 Level
ERB	ethical review board
ET	early termination
FAS	full analysis set
FDA	Food and Drug Administration
F-DBL	final database lock
FSH	follicle-stimulating hormone
fXa	clotting factor Xa
GBD	Global Burden of Disease
GCP	good clinical practice
GGT	gamma-glutamyltransferase

GPS	Global Patient Safety
HADS	Hospital Anxiety and Depression Scale
HBcAb	hepatitis B core antibody
HBsAb	hepatitis B surface antibody
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
IAS	interim analysis set
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation
IFNγ	interferon gamma
Ig	immunoglobulin
IL	interleukin
INR	international normalized ratio
IP	investigational product: a pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
IRB	institutional review board
IV	intravenous
IWRS	interactive web-response system
JAK	Janus kinase
LS	least squares
MACE	major adverse cardiovascular events
MCMC	Markov chain Monte Carlo method
MCS	mental component score

MedDRA	Medical Dictionary for Regulatory Activities
MHC	major histocompatibility complex
mFAS	modified full analysis set
MI	multiple imputation
mLOCF	modified last observation carried forward
MMRM	mixed-effects model of repeated measures
mRNA	messenger RNA
MTHFR	methylene tetrahydrofolate reductase
NRI	nonresponder imputation
PCS	physical component score
PD	pharmacodynamics(s)
PE	pulmonary embolism
PK	pharmacokinetic(s)
pMI	placebo multiple imputation
PO-DBL	primary outcome database lock
PPD	purified protein derivative
PPS	per-protocol set
PRO/ePRO	patient-reported outcome/electronic patient-reported outcome
PT	prothrombin time
PTT	partial thromboplastin time
QD	once daily administration
RA	rheumatoid arthritis
RBC	red blood cell
RNA	ribonucleic acid
SAE	serious adverse event
SALT	Severity of Alopecia Tool
$SALT_{30}$	at least 30% improvement from Baseline in SALT score

$SALT_{50}$	at least 50% improvement from Baseline in SALT score
$SALT_{75}$	at least 75% improvement from Baseline in SALT score
$SALT_{90}$	at least 90% improvement from Baseline in SALT score
SAP	statistical analysis plan
SF-36	Short-Form Health Survey
SOA	Schedule of Events
STAT	signal transducer and activator of transcription
SUSAR	suspected unexpected serious adverse reaction
TB	tuberculosis
TBL	total bilirubin level
TEAE	treatment-emergent adverse event: an untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, which does not necessarily have to have a causal relationship with this treatment.
TSH	thyroid-stimulating hormone
TYK2	tyrosine kinase 2
ULN	upper limit of normal
VAS	Visual Analog Scale
VTE	venous thromboembolic event
WBC	white blood cell

Appendix 2. Clinical Laboratory Tests

Hematology[a,b]	**Clinical Chemistry**[a,b]
Hemoglobin	**Serum Concentrations of:**
Hematocrit	Sodium
Erythrocyte count (RBC)	Potassium
Absolute Reticulocyte Count	Total bilirubin
Mean cell volume	Direct bilirubin
Mean cell hemoglobin	Alkaline phosphatase
Mean cell hemoglobin concentration	Alanine aminotransferase (ALT)
Leukocytes (WBC)	Aspartate aminotransferase (AST)
Platelets	Blood urea nitrogen (BUN)
Absolute counts of:	Creatinine
Neutrophils, segmented	Cystatin C
Neutrophils, juvenile (bands)	Uric acid
Lymphocytes	Calcium
Monocytes	Glucose
Eosinophils	Albumin
Basophils	Total protein
	Estimated glomerular filtration rate (eGFR)[e]
	Creatine phosphokinase (CPK)
Urinalysis[a,b,d]	
Color	**Other Tests**[a]
Specific gravity	Hepatitis B Surface antigen (HBsAg)[f]
pH	Anti-Hepatitis B Core antibody (HBcAb)[f]
Protein	Hepatitis B virus (HBV) DNA[l]
Glucose	Anti-Hepatitis B Surface antibody (HBsAb)[f]
Ketones	Human immunodeficiency virus (HIV)[f]
Bilirubin	Hepatitis C antibody[f,g]
Urobilinogen	Thyroid-stimulating hormone (TSH)[f]
Blood	Exploratory storage samples (serum, plasma and mRNA)
Leukocyte esterase	Pregnancy Test[h]
Nitrite	Follicle-stimulating hormone[f,i]
	QuantiFERON-TB Gold[j] (central testing preferred) or T-SPOT.*TB*[k] (local testing)
Lipids[a,c]	PPD (local testing/reading)
Total cholesterol	Baricitinib plasma levels
Low-density lipoprotein	Serum immunoglobulin (IgE)
High-density lipoprotein	
Triglycerides	

Abbreviations: DNA = deoxyribonucleic acid; mRNA = messenger ribonucleic acid; FSH = follicle-stimulating hormone; HBV = hepatitis B virus; IgE = immunoglobulin E; PPD = purified protein derivative; RBC = red blood cell; TB = tuberculosis; WBC = white blood cell.

[a] Assayed by Sponsor-designated laboratory.

[b] Unscheduled or repeat blood chemistry, hematology, and urinalysis panels may be performed at the discretion of the investigator, as needed.

[c] Fasting lipid profile. Patients should not eat or drink anything except water for 8 hours prior to test. If a patient attends these visits in a nonfasting state, this will not be considered a protocol violation.

[d] Microscopic examination of sediment performed only if abnormalities are noted on the routine urinalysis.

[e] Estimated glomerular filtration rate for serum creatinine calculated by the central laboratory using the Chronic Kidney Disease Epidemiology Collaboration creatinine 2009 equation.

[f] Test required at Visit 1 only to determine eligibility of patient for the study.

[g] A positive hepatitis C antibody result will be confirmed with an alternate hepatitis C method.

[h] For all women of childbearing potential, a serum pregnancy test will be performed at Visit 1 and a local urine pregnancy test will be performed at Visit 2 and at all subsequent study visits. If required per local regulations and/or institutional guidelines, pregnancy testing can occur at other times during the study treatment period.

[i] To confirm postmenopausal status for women ≥40 and <60 years of age who have had a cessation of menses, an FSH test will be performed. Nonchildbearing potential is defined as an FSH ≥40 mIU/mL and a cessation of menses for at least 12 months.

[j] The QuantiFERON-TB Gold test (central testing) is the preferred alternative to the PPD test for the evaluation of TB infection, and it may be used instead of the PPD test or T-SPOT.*TB* test. If the QuantiFERON-TB Gold test is indeterminate, 1 retest is allowed.

[k] T-SPOT.*TB* must be read locally.

[l] HBV DNA testing will be done in those patients who are HBcAb+ at screening, regardless of HBsAb status.

Appendix 3. Study Governance Considerations

Appendix 3.1. Regulatory and Ethical Considerations, Including the Informed Consent Process

Appendix 3.1.1. Informed Consent

The investigator is responsible for:

- ensuring that the patient understands the nature of the study, the potential risks and benefits of participating in the study, and that their participation is voluntary.
- ensuring that informed consent is given by each patient or legal representative. This includes obtaining the appropriate signatures and dates on the informed consent form (ICF) prior to the performance of any protocol procedures and prior to the administration of investigational product.
- answering any questions the patient may have throughout the study and sharing in a timely manner any new information that may be relevant to the patient's willingness to continue his/her participation in the study.
- ensuring that a copy of the ICF is provided to the participant or the participant's legal representative and is kept on file.
- ensuring that the medical record includes a statement that written informed consent was obtained before the participant was enrolled in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Appendix 3.1.2. Recruitment

Lilly is responsible for the central recruitment strategy for patients. Individual investigators may have additional local requirements or processes.

Appendix 3.1.3. Ethical Review

The investigator must give assurance that the ethical review board (ERB) was properly constituted and convened, as required by International Council for Harmonisation (ICH) guidelines and other applicable laws and regulations.

Documentation of ERB approval of the protocol and the ICF must be provided to Lilly before the study may begin at the investigative site(s). Lilly or its representatives must approve the ICF, including any changes made by the ERBs, before it is used at the investigative site(s). All ICFs must be compliant with the ICH guideline on Good Clinical Practice (GCP).

The study site's ERB(s) should be provided with the following:

- the protocol and related amendments and addenda, current Investigator's Brochure (IB) and updates during the course of the study
- informed consent form
- other relevant documents (eg, curricula vitae, advertisements)

Appendix 3.1.4. Regulatory Considerations

This study will be conducted in accordance with the protocol and with the:

- consensus ethics principles derived from international ethics guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences International Ethical Guidelines
- applicable ICH GCP Guidelines
- applicable laws and regulations

Some of the obligations of the Sponsor will be assigned to a third party.

Appendix 3.1.5. Investigator Information

Physicians with expertise in the diagnosis and treatment of alopecia areata will participate as investigators in this clinical trial.

Appendix 3.1.6. Protocol Signatures

The Sponsor's responsible medical officer will approve the protocol, confirming that, to the best of his/her knowledge, the protocol accurately describes the planned design and conduct of the study.

After reading the protocol, each principal investigator will sign the protocol signature page and send a copy of the signed page to a Lilly representative.

Appendix 3.1.7. Final Report Signature

Lilly will select a qualified investigator(s) from among investigators participating in the design, conduct, and/or analysis of the study to serve as the clinical study report (CSR) coordinating investigator. If this investigator is unable to fulfill this function, another investigator will be chosen by Lilly to serve as the CSR coordinating investigator.

The CSR coordinating investigator will sign the final CSR for this study, indicating agreement that, to the best of his/her knowledge, the report accurately describes the conduct and results of the study.

The Sponsor's responsible medical officer and statistician will approve the final CSR for this study, confirming that, to the best of his/her knowledge, the report accurately describes the conduct and results of the study.

Appendix 3.2. Data Quality Assurance

To ensure accurate, complete, and reliable data, Lilly or its representatives will do the following:

- provide instructional material to the study sites, as appropriate.
- sponsor start-up training to instruct the investigators and study coordinators. This training will give instruction on the protocol, the completion of the case report forms (CRFs), and study procedures.

- make periodic visits to the study site.
- be available for consultation and stay in contact with the study site personnel by mail, telephone, and/or fax.
- review and evaluate CRF data and use standard computer edits to detect errors in data collection.
- conduct a quality review of the database.

In addition, Lilly or its representatives will periodically check a sample of the patient data recorded against source documents at the study site. The study may be audited by Lilly or its representatives, and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

The investigator will keep records of all original source data. This might include laboratory tests, medical records, and clinical notes. If requested, the investigator will provide the Sponsor, applicable regulatory agencies, and applicable ERBs with direct access to original source documents.

Appendix 3.2.1. Data Capture System

An electronic data capture system will be used in this study. The site maintains a separate source for the data entered by the site into the Sponsor-provided electronic data capture system.

Electronic patient-reported outcome (ePRO) measures (eg, a rating scale) and electronic clinical outcome assessments (eCOAs) are entered into an ePRO/eCOA instrument at the time that the information is obtained. In those instances where there is no prior written or electronic source data at the site, the ePRO/eCOA instrument record will serve as the source.

If ePRO/eCOA records are stored at a third-party site, investigator sites will have continuous access to the source documents during the study and will receive an archival copy at the end of the study for retention.

Any data for which the ePRO/eCOA instrument record will serve to collect source data will be identified and documented by each site in that site's study file.

Case report form data will be encoded and stored in InForm.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system. Data will subsequently be transferred from the central vendor to the Lilly data warehouse.

Data from complaint forms submitted to Lilly will be encoded and stored in the global product complaint management system.

Appendix 3.3. Study and Site Closure

Appendix 3.3.1. Discontinuation of Study Sites

Study site participation may be discontinued if Lilly, the investigator, or the ERB of the study site judges it necessary for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and GCP.

Appendix 3.3.2. Discontinuation of the Study

The study will be discontinued if Lilly judges it necessary for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and GCP. Study termination may occur in a specific country or region when baricitinib is approved for the treatment of AA and becomes reimbursed or commercially available in that country or region, or a negative regulatory action or opinion is received in that country or region.

Appendix 3.4. Publication Policy

The publication policy for Study I4V-MC-JAHO is described in the Clinical Trial Agreement.

Appendix 4. Hepatic Monitoring Tests for Treatment-Emergent Abnormality

Selected tests may be obtained in the event of a treatment-emergent hepatic abnormality and may be required in follow-up with patients in consultation with the Lilly, its designee, or the clinical research physician.

Hepatic Monitoring Tests

Hepatic Hematology[a]	**Haptoglobin[a]**
Hemoglobin	
Hematocrit	**Hepatic Coagulation[a]**
Erythrocyte count (RBC)	Prothrombin Time
Leukocytes (WBC)	Prothrombin Time, INR
Neutrophils, segmented	
Lymphocytes	**Hepatic Serologies[a,b]**
Monocytes	Hepatitis A antibody, total
Eosinophils	Hepatitis A antibody, IgM
Basophils	Hepatitis B surface antigen
Platelets	Hepatitis B surface antibody
	Hepatitis B Core antibody
Hepatic Chemistry[a]	Hepatitis C antibody
Total bilirubin	Hepatitis E antibody, IgG
Direct bilirubin	Hepatitis E antibody, IgM
Alkaline phosphatase	
ALT	**Anti-nuclear antibody[a]**
AST	
GGT	**Alkaline Phosphatase Isoenzymes[a]**
CPK	
	Anti-smooth muscle antibody (or anti-actin antibody)[a]

Abbreviations: ALT = alanine aminotransferase; AST = aspirate aminotransferase; CPK = creatinine phosphokinase; GGT = gamma-glutamyl transferase; Ig = immunoglobulin; INR = international normalized ratio; RBC = red blood cell; WBC = white blood cell.

a Assayed by Lilly-designated or local laboratory.

b Reflex/confirmation dependent on regulatory requirements and/or testing availability.

Appendix 5. Liver Function Testing and Hepatic Safety Monitoring

Liver Function Testing and Hepatic Safety Monitoring

Analyte	Exclusion Criteria	Additional Hepatic Testing	Hepatic eCRF Reporting	Temporary Interruption of IP	Permanent Discontinuation of IP after Consultation with the Lilly-Designated Medical Monitor
Protocol Section	Section 6.2	Section 9.4.10.1	Section 9.4.10.1	Section 8.1.1	Section 8.1.1
ALT/AST	≥2 × ULN	As per protocol only ALT >3 × ULN	As per protocol only ALT ≥5 × ULN on ≥2 consecutive tests	≥5 × ULN	• >8 × ULN • >5 × ULN for >2 weeks • >3 × ULN **AND** TBL >2 × ULN or INR >1.5 • >3 × ULN with symptoms[a]
ALP	≥2 × ULN	≥2 × ULN	≥2 × ULN on ≥2 consecutive tests	No applicable criteria	• >3 × ULN • >2.5 × ULN **AND** TBL >2 × ULN • >2.5 × ULN with symptoms[a]
TBL	≥1.5 × ULN	≥2 × ULN	≥2 × ULN (excluding Gilbert's syndrome)	No applicable criteria	• ALT or AST (as per protocol) >3 × ULN **AND** TBL >2 × ULN • ALP >2.5 × ULN **AND** TBL >2 × ULN

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; eCRF = electronic case report form; INR = international normalized ratio; IP = investigational product; TBL = total bilirubin level; ULN = upper level of normal.

[a] Fatigue, nausea, vomiting, right upper-quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%).

Appendix 6. Additional Information on Statistical Methods

Appendix 6.1. Statistical Methods for Additional Analyses

Appendix 6.1.1. Placebo Multiple Imputation

The placebo multiple imputation (pMI) method may be used as an additional analysis for the analysis of the primary efficacy endpoint, as well as key secondary endpoints. Multiple imputations are used to replace missing outcomes (defined in SAP) (Severity of Alopecia Tool [SALT] score, patient-reported outcome [PRO] for Scalp Hair Assessment score, etc.) for drug- and placebo-treated patients using multiple draws from the posterior predictive distribution estimated from the placebo arm. The binary outcomes (SALT ≤20, PRO for Scalp Hair Assessment score of 0 or 1 with a ≥2-point improvement, etc.) will be derived from the imputed data.

Data are processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcomes at visits t=1,.., T.

1. *Initialization:* Set t=0 (Baseline visit)
2. *Iteration:* Set t=t+1. Create a data set combining records from drug- and placebo-treated patients with columns for covariates **X** and outcomes at visits 1,..,t with outcomes for all drug-treated patients set to missing at visit t and set to observed or imputed values at visits 1,..,t-1.
3. *Imputation:* Run Bayesian regression in SAS® PROC MI on this data to impute missing values for visit t using previous outcomes for visits 1 to t-1 and baseline covariates. Note that only placebo data will be used to estimate the imputation model because no outcome is available for drug-treated patients at visit t.
4. Replace imputed data for all drug-treated patients at visit t with their observed values, whenever available up to permanent study drug discontinuation or study discontinuation. If t < T, then go to Step 2, otherwise proceed to Step 5.
5. Repeat steps 1-4, m times with different seed values to create m imputed complete data sets.

Note that a Markov chain Monte Carlo method will be used to create a monotone missing pattern first in the presence of intermittent missing data.

Analysis For the primary and key secondary efficacy endpoints, the missing outcomes will be derived from the imputed data for each patient before fitting into the analysis model. A logistic regression or analysis of covariance (ANCOVA)will be applied, as appropriate.

The number of imputed data sets will be m=100 and a 6-digit seed value will be pre-specified for each analysis. Within the program, the seed will be used to generate the m seeds needed for imputation. The initial seed values are given below:

Analysis	Seed Value
Proportion of patients achieving SALT ≤20 at Weeks 16, 24, and 36	123450
Proportion of patients achieving a PRO for Scalp Hair Assessment 0 or 1 with a ≥2-point improvement from Baseline at Week 36 among patients with a score of ≥3 at Baseline	123451
Proportion of patients achieving an absolute SALT score ≤ 10 at Weeks 24 and 36	123450
Proportion of patients achieving $SALT_{90}$ at Week 36	123450
Proportion of patients achieving $SALT_{50}$ at Week 12	123450
Percent change from Baseline in SALT score at Week 36	123450
Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week 36 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline)	123452
Proportion of patients achieving ClinRO Measure for EL Hair Loss 0 or 1 with ≥2-point improvement from Baseline at Week36 (among patients with ClinRO Measure for EL Hair Loss ≥2 at Baseline)	123453

Abbreviations: ClinRO = clinician-reported outcome; EB = eyebrow; EL = eyelash; PRO = patient-reported outcome; SALT = Severity of Alopecia Tool; $SALT_{50}$ = at least 50% improvement from Baseline in SALT score; $SALT_{90}$ = at least 90% improvement from Baseline in SALT score.

The final inference on treatment difference is conducted from the multiple data sets using Rubin's combining rules, as implemented in SAS® PROC MIANALYZE.

Thus, in the effectiveness context, pMI assumes no pharmacological benefit of the drug after dropout. It limits bias by taking into account study/placebo effects. In the efficacy context, pMI is a specific form of a missing not at random analysis expected to yield a conservative estimate of efficacy.

Appendix 6.1.2. Tipping Point Analyses

To investigate the missing data mechanism, an additional analysis using multiple imputation (MI) under the missing not at random assumption may be provided for the primary objective, which compares the SALT ≤20 response rate of each of the baricitinib 4-mg and 2-mg doses to placebo at Week 36. The tipping point analysis may also be used as an additional sensitivity analysis for some key secondary objectives.

All patients in the full analysis set (FAS) population are included. Within each analysis, the most extreme case will be considered, in which all missing data (defined in SAP) for patients randomized to the baricitinib doses will be imputed using the worst possible result, and all missing data (defined in SAP) for patients randomized to placebo will be imputed with the best possible result. Treatment differences will be analyzed using logistic regression or analysis of covariance (ANCOVA), as appropriate.

For continuous variables, the following process will be used to determine the tipping point:

1. To handle intermittent missing visit data, a Markov chain Monte Carlo method (SAS® Proc MI with MCMC option) will be used to create a monotone missing pattern.
2. A set of Bayesian regressions (using SAS® Proc MI with MONOTONE option) will be used for the imputation of monotone dropouts. Starting from the first visit with at least 1 missing value, the regression models will be fit sequentially with treatment as a fixed effect and values from the previous visits as covariates.
3. A delta score is added to all imputed scores at the timepoint where the analysis is conducted for patients in the baricitinib treatment groups, thus, worsening the imputed value. The delta score is capped for patients, based on the range of the outcome measure being analyzed.
4. Treatment differences between baricitinib and placebo are analyzed for each imputed data set using ANCOVA. Results across the imputed data sets are aggregated using SAS® Proc MIANALYZE in order to compute a p-value for the treatment comparisons for the given delta value.
5. Steps 3 and 4 are repeated, and the delta value added to the imputed baricitinib scores is gradually increased. The tipping point is identified as the delta value which leads to a loss of statistical significance (aggregated p-value >0.05) when evaluating baricitinib relative to the placebo group.

As a reference, for each delta value used in Steps 3-5, a fixed selection of delta values (ranging from slightly negative to slightly positive) will be added to imputed values in the placebo group, and Step 4 will be performed for the combination. This will result in a 2-d table for each timepoint of interest, with the columns representing the delta values added to the imputed placebo responses, and the rows representing the delta values added to the imputed baricitinib responses. Separate 2-d tables will compare each baricitinib dose group to placebo.

A similar process will be used for the categorical variables:

1. Missing responses in the baricitinib groups will be imputed with a range of low response probabilities, including probabilities of 0, 0.1, and 0.2.
2. For missing responses in the placebo group, a range of response probabilities (eg, probability = 0, 0.2, … 1) will be used to impute the missing values. Multiple imputed data sets will be generated for each response probability.
3. Treatment differences between baricitinib and placebo are analyzed for each imputed data set using logistic regression. Results across the imputed data sets are aggregated using SAS® Proc MIANALYZE in order to compute a p-value for the treatment comparisons for the given response probability. If the probability values do not allow for any variation between the multiple imputed data sets (eg, all missing responses in the

placebo and baricitinib groups are imputed as responders and nonresponders, respectively), then the p-value from the single imputed data set will be used.

The tipping point is identified as the response probability value within the placebo group that leads to a loss of statistical significance when evaluating baricitinib relative to placebo.

For tipping point analyses the number of imputed data sets will be m=100 and the seed value to start the pseudorandom number generator of SAS Proc MI (same values for MCMC option and for MONOTONE option) will be:

Analysis	**Seed value**
Proportion of patients achieving SALT ≤20 at Week 36	123461

Appendix 6.2. Graphical Testing Procedures

The primary hypotheses $H_{1,1,0}$ and $H_{2,1,0}$ will be first tested at a two-sided α=0.04 and 0.01, respectively. If neither of the null hypotheses is rejected, no further testing is conducted, as the α for that test is considered "spent" and cannot be passed to other endpoints. If at least one of the null hypotheses is rejected, the testing process continues, with the remaining α propagated according to the weights on the corresponding edges displayed in (Figure JAHO.3). The testing process continues as long as there is at least one hypothesis in the scheme that can be rejected at its allocated α level at that point. Each time a hypothesis is rejected, the graph is updated to reflect the reallocation of α, which is considered "recycled" by Alosh et al. (2014). This iterative process of updating the graph and reallocating α is repeated until all hypotheses have been tested or when no remaining hypotheses can be rejected at their corresponding α levels.

There will be no adjustment for multiple comparisons for any other analyses.

Appendix 7. Monitoring Tests for Confirmed VTE

Selected tests may be obtained in the event of a confirmed venous thromboembolic event (VTE) and may be required in follow-up with patients in consultation with Eli Lilly and Company, its designee, or the clinical research physician.

Protein C Functional
Protein S Clottable
Antithrombin III
APC Resistance
PT
APTT
Fibrinogen
Cardiolipin Antibodies
PT Gene
Factor VIII C Assay
Hexagonal Phase Phospholipid Neutralization
C-Reactive Protein
PTT Incubated Mixing
Dilute Russell Viper Venom
Platelet Neutralization
Factor V Leiden
MTHFR
Thrombin Time
Reptilase
Fibrinogen Antigen
Protein C Immunologic
Protein S Immunologic
Heparin fXa Inhibition

Abbreviations: APC = activated protein C; APTT = activated partial thromboplastin time; fXa = clotting factor Xa; MTHFR = methylene tetrahydrofolate reductase; PT = prothrombin time; PTT = partial thromboplastin time.

Appendix 8. Provisions for Changes in Study Conduct During Exceptional Circumstances

Implementation of this appendix

The changes to procedures described in this appendix are temporary measures intended to be used only during specific time periods as directed by the sponsor in partnership with the investigator.

Exceptional circumstances

Exceptional circumstances are rare events that may cause disruptions to the conduct of the study. Examples include pandemics or natural disasters. These disruptions may limit the ability of the investigators, participants, or both to attend on-site visits or to conduct planned study procedures.

Implementing changes under exceptional circumstances

In an exceptional circumstance, after receiving the sponsor's written approval, sites may implement changes if permitted by local regulations.

After approval by local ERBs, regulatory bodies and any other relevant local authorities, implementation of these exceptional circumstances changes will not typically require additional notification to these groups, unless they have specific conditions in which notification is required. To protect the safety of study participants, urgent changes may be implemented before approval but need to be reported as soon as possible. All approvals must be retained in the study records.

In the event written approval is granted by the sponsor for changes in study conduct, additional written guidance, if needed, will be provided by the sponsor.

Considerations for making a change

The prevailing consideration for making a change is ensuring the safety of study participants. Additional important considerations for making a change are compliance with Good Clinical Practice, enabling participants to continue safely in the study and maintaining the integrity of the study.

Informed Consent

The site should document the participant's verbal consent for having remote visits and remote dispensing of IP, ancillaries, prior to implementation of these activities.

Additional consent from the participant will be obtained, if required, for:

- participation in remote visits, as defined in Section "Remote Visits,"
- a change in the method, location, or both, of study intervention administration,
- dispensation of additional study intervention during an extended treatment period,
- alternate delivery of study intervention and ancillary supplies, and

- provision of their personal or medical information required prior to implementation of these activities.

Changes in Study Conduct During Exceptional circumstances

Changes in study conduct not described in this appendix, or not consistent with applicable local regulations, are not allowed.

The following changes in study conduct will not be considered protocol deviations.

1. **Remote visits**

   Visit 8 (Week 36) requires a live clinical assessment because it is the primary outcome visit. To facilitate an onsite visit, the visit window may be extended (see below in Adjustment to Visit Windows of this appendix).

   All other visits (including assessment of loss of benefit during randomized down titration post Week 52), **may be conducted remotely** if the visit could not otherwise be conducted due to local or national restrictions.

In source documents and the CRF, the study site should capture the visit location and method, with a specific explanation for any data missing because of missed in-person site visits.

Telemedicine: Live, onsite clinical efficacy assessments are preferred for ALL visits; however, if a remote visit is the only option, telephone or technology-assisted virtual visits, or both, are acceptable to complete appropriate assessments (with the exception of Visit 8). Assessments to be completed in this manner include:

- PRO Scalp, SALT Score, and ClinROs for Eyebrows/Eyelashes if Visual Assessment by telemedicine is possible. If NO visual assessment AND ONLY verbal assessment by phone is possible. SALT, ClinROs for Eyelash/Eyelash, Nails, PROs for Eye Irritation should not be assessed.
- AEs and product complaints
- concomitant medications
- C-SSRS (Since Last Visit Version), Self-Harm Supplement Form, and Self-Harm
- Photos should be captured at next onsite visit after a remote V8 or V11/ET

Regardless of the type of remote visits implemented, the protocol requirements regarding the reporting of adverse events (AEs), serious adverse events (SAEs), and product complaints remain unchanged. Furthermore, every effort should be made to enable participants to return to on-site visits as soon as reasonably possible, while ensuring the safety of both the participants and the site staff.

2. **Local laboratory testing option**

   Local laboratory testing may be conducted in lieu of central laboratory testing. The local laboratory must be qualified in accordance with applicable local regulations.

   - Obtain local labs for safety (hematology, chemistry) and urine pregnancy when applicable, as per the study protocol schedule of activities.

- A urine pregnancy test should be sent to female patients of child bearing potential. The patient should conduct the test and provide verification of results to the site per a method agreed to between the patient and the site. Investigational product must be temporarily held until site can review photo of pregnancy test results.
- All labs will be reviewed by the investigators. Lilly Medical should be informed of any labs that meet criteria for temporary or permanent study drug discontinuation.
- Sign and date review of local labs per normal process and follow-up with the patient as needed. Results will not be recorded in the eCRF.
- Safety labs should be obtained at a minimum of every 12 weeks and IP should not be dispensed until these can be collected and reviewed.

3. Study intervention and ancillary supplies

When a participant is unable to go to the site to receive study supplies during normal on-site visits, the site should work with the sponsor to determine appropriate actions. These actions may include:

- asking the participant to go to the site and receive study supplies from site staff without completion of a full study visit,
- asking the participant's designee to go to the site and receive study supplies on a participant's behalf,
- arranging delivery of study supplies.

These requirements must be met before action is taken:

- Sponsor approves the alternative method of delivery, taking local regulatory requirements into consideration.
- Participant consents verbally to alternate method of delivery.
- Site confirms the participant's receipt of the trial supplies.
- Site/sponsor confirms appropriate ethics review board notification.
- Alternate delivery of IP should be performed in a manner that does not compromise treatment blinding and ensures product integrity. The existing protocol requirements for product accountability remain unchanged, including verification of participant's receipt of study supplies.
- When delivering supplies to a location other than the study site (for example, participant's home), the investigator, sponsor, or both should ensure oversight of the shipping process to ensure accountability and product quality (that is, storage conditions maintained and intact packaging upon receipt).
- Instructions may be provided to the participant or designee on the final disposition of any unused or completed study supplies.

4. Screening period guidance

Not Applicable

5. Adjustments to Visit Windows

Whenever possible and safe to do so, as determined by the investigator's discretion, participants should complete the usual Schedule of Activities. To maximize the possibility that these visits can be conducted as on-site visits, the windows for visits may be adjusted, upon further guidance from the sponsor. This minimizes missing data and preserves the intended conduct of the study.

This table describes the allowed adjustments to visit windows.

Visit Number	Tolerance
Visits 8	Visits 8 (Week 36) may be conducted within 7 days before the intended date, or up to 28 days after the intended date
All other Visits	All other visits should be conducted as per the visit window the stated in the Schedule of Activities in Section 2

Documentation

Changes to study conduct will be documented:

- Sites will identify and document the details of how participants, visit types, and conducted activities were affected by exceptional circumstances. Dispensing/shipment records of study intervention and relevant communications, including delegation, should be filed with site study records.
- Source documents generated at a location other than the study site should be part of the investigator's source documentation and should be transferred to the site in a secure and timely manner.

Appendix 9. Protocol Amendment I4V-MC-JAHO(e) Summary

Overview

Protocol I4V-MC-JAHO A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Operationally Seamless, Adaptive Phase 2/3 Study to Evaluate the Efficacy and Safety of Baricitinib in Adult Patients with Severe or Very Severe Alopecia Areata (BRAVE-AA1) has been amended. The new protocol is indicated by amendment (e) and will be used to conduct the study in place of any preceding version of the protocol.

This protocol has been amended to:

- finalize key secondary endpoints prior to primary outcome database lock,
- update and clarify statistical analysis,
- modify graphical testing to align with the updates to key secondary endpoints,
- add Appendix 8 to allow for remote visits in the event of extraordinary events, and
- correct typographical errors, clarify study procedures, and increase flexibility of overall study.

The overall changes and rationale for the changes made to this protocol are described in the following table.

Amendment Summary for Protocol I4V-MC-JAHO Amendment (e)

Section # and Name	Description of Change	Brief Rationale
Section 1. Synopsis	• Added proportion of patients achieving $SALT_{50}$ at Week 12 to Key Secondary Objectives • Moved Week 24 timepoint for Proportion of patient achieving a $SALT_{90}$ to Other Secondary Endpoints • Changed timepoint for percent change from baseline in SALT score to Week 36 and moved Week 12 to Other Secondary Endpoints • Moved proportion of patients achieving $SALT_{50}$ at Week 12 from Other Secondary Endpoints to Key Secondary Endpoints • Moved ClinROs for EB and EL to Key Secondary Endpoints at Week 36 from Other Secondary Endpoints • Moved Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline) from Other Secondary Endpoints to Key Secondary Endpoints • Moved Weeks 16 and 24 for Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 to Other Secondary Endpoints • Clarified ClinRO and PRO endpoints • Change from Baseline in SALT score at Weeks 12, 16, 24, and 36 added to Other Secondary • Added clarifying statement in regard to the potential for additional database locks and Interim Analysis • Clarified Mean change from Baseline in HADS total score at Weeks 24 and 36 to specify HADS-A and HADS-D • Changed 95% CI will be reported, to may be reported • Removed Baseline testing for Fisher exact test	• Updated key secondary endpoints to reflect the analysis that will be conducted • For flexibility • Error
Section 2. SoA	• Added leading statement regarding addition of Appendix 8 which addresses provisions in study conduct under exceptional circumstances	• For clarity on addition of Appendix 8
Section 4. Objectives and Endpoints	• Added proportion of patients achieving $SALT_{50}$ at Week 12 to Key Secondary Objectives • Moved Week 24 timepoint for Proportion of patients achieving a $SALT_{90}$ to Other Secondary Endpoints • Changed timepoint for percent change from baseline in SALT score to Week 36 and moved Week 12 to Other Secondary Endpoints	• Updated key secondary endpoints to reflect the analysis that will be

	• Moved proportion of patients achieving $SALT_{50}$ at Week 12 from Other Secondary Endpoints to Key Secondary Endpoints • Moved ClinROs for EB and EL to Key Secondary Endpoints at Week 36 from Other Secondary Endpoints • Moved Proportion of patients achieving ClinRO Measure for EB Hair Loss 0 or 1 with ≥2-point improvement from baseline at Week 36 (among patients with ClinRO Measure for EB Hair Loss ≥2 at Baseline) from Other Secondary Endpoints to Key Secondary Endpoints • Moved Weeks 16 and 24 for Proportion of patients with PRO for Scalp Hair Assessment score of 0 or 1 to Other Secondary Endpoints • Clarified ClinRO and PRO endpoints • Change from Baseline in SALT score at Weeks 12, 16, 24, and 36 added to Other Secondary • Added clarifying statement in regard to the potential for additional database locks and Interim Analysis • Clarified Mean change from Baseline in HADS total score at Weeks 24 and 36 to specify HADS-A and HADs-D • Added EQ-5D-5L and SF-36 abbreviations	conducted
Section 5.1.1. Stage 1 (Includes Phase 2 Portion and Phase 3 Portion before Decision Point) Section 5.1.3. Stage 2 (Remaining Portion of Phase 3 Study after the Decision Point)	• Removed wording referring to AA severity to improve clarity of stratification	• For clarity based on regulatory feedback
Section 6.1. Inclusion Criteria	• Updated pregnancy language	• For Clarity
Section 7.2. Method of Treatment Assignment	• Removed wording referring to AA severity to improve clarity of stratification • Visit and Week information corrected	• Regulatory feedback • Error
Section 9.4.4. Laboratory Tests	• Clarified blood and urine collection for home visits are permitted under normal protocol operations	• For flexibility in study conduct
Section 10.3.1. General Statistical Considerations	• Sentence added to allow for additional database locks and interim analyses being conducted	• For flexibility and to support scientific disclosures
Section 10.3.1.1. Analysis Methods	• 95% CI will also be reported changed to may be reported	• Error

Section 10.3.1.2. Missing Data Imputation	• Statements removed to clarify sensitivity analysis may be conducted • The specification for additional analysis was updated • pMI analysis was changed from will be included to may be included	• For flexibility
Section 10.3.1.3. Multiple Comparisons/Multiplicity	• Update Graphical testing procedures to reflect planned analysis (Figure JAHO.3) • Key Secondaries Null Hypotheses Updated to reflect updates to Key Secondary Objectives	• Updated based on revision to Key Secondary Endpoints
Section 10.3.3.1. Primary Analyses	• Statement updated to clarify that additional analysis of the primary efficacy outcome may be performed	• Updated based on Regulatory feedback
Section 10.3.3.2. Secondary Analyses	• Statement updated to clarify that additional analysis of the secondary efficacy outcome may be performed	• Updated based on Regulatory feedback
Section 10.3.4. Safety Analyses	• Updated wording to clarify data included in the safety analysis	• For clarity
Section 10.3.6.1. Subgroup Analyses	• Added baseline value wording to logistic model	• For clarity
Section 10.3.7.3.2. Week 36 Primary Outcome Analysis and Other Regulatory Submission Activities	• Updated 4 month safety update database lock to 3 to 4 month safety update data base lock • Added additional efficacy/safety interim analyses are also done to support scientific disclosures.	• For flexibility
Appendix 6.1.1. Placebo Multiple Imputation	• Sentence added in stating that a Markov chain Monte Carlo method will be used • Analysis table adjusted to reflect planned analysis and updated objectives • Removed statement about mLOCF	• For clarity
Appendix 8 Provisions for Changes in Study Conduct During Exceptional Circumstances	• Added Provisions for Changes in Study Conduct During Exceptional Circumstances Appendix	• Allows for continued conduct of the study using remote visits and increases flexibility

Abbreviations: AA = alopecia areata; CI = Confidence Interval; ClinRO = clinician-reported outcome; EB = eyebrow; EL = eyelash; EQ-5D-5L= European Quality of Life – 5 Dimensions – 5 Level; mLOCF = modified last observation carried forward; pMI = placebo multiple imputation; PRO = patient-reported outcome; SALT = Severity of Alopecia Tool; $SALT_{50}$ = at least 50% improvement from Baseline in Severity of Alopecia Tool score; $SALT_{90}$ = at least 90% improvement from Baseline in Severity of Alopecia Tool score; SF-36 = Short Form-36 Health Survey acute version 2.

Leo Document ID = fa08c39c-2a1c-4e5f-9d93-23441e9d7e8f

Approver: PPD

Approval Date & Time: 02-Nov-2020 19:28:21 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 02-Nov-2020 20:59:54 GMT

Signature meaning: Approved