CLINICAL STUDY REPORT SYNOPSIS

Sponsor: Pfizer, Inc.

Investigational Product: PF-04965842 (abrocitinib)

Clinical Study Report Synopsis: Protocol B7451013

Protocol Title: A Phase 3 Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Multi-Center Study to Evaluate the Efficacy and Safety of PF-04965842 Monotherapy in Subjects Aged 12 Years and Older, With Moderate to Severe Atopic Dermatitis

Investigators: Refer to Appendix 16.1.4.1 for a list of investigators involved in this study.

Study Center(s): A total of 391 subjects were randomized at 102 sites in 13 countries, including sites in the United States (n=19), Poland (n=14), Republic of Korea (n=10), Japan (n=8), Australia (n=7), Bulgaria (n=7), Canada (n=7), Germany (n=7), United Kingdom (n=6), China (n=5), Latvia (n=5), Hungary (n=4), and Czech Republic (n=3). Refer to Appendix 16.1.4.1 for a list of sites involved in this study.

Publications Based on the Study: None

Study Initiation Date: First Subject First Visit (FSFV): 29 June 2018.

Study Completion Date: Last Subject Last Visit (LSLV): 13 August 2019.

Report Date: 13 January 2020

Previous Report Date(s): Not Applicable.

Phase of Development: Phase 3

Primary and Secondary Study Objectives and Endpoints:

Table S1. Study Objectives and Endpoints

Type	Objective	Endpoint
Primary		
Efficacy	• To assess the efficacy of PF-04965842 compared with placebo in subjects aged 12 years and older with moderate to severe atopic dermatitis (AD).	• Response based on the Investigator's Global Assessment (IGA) score of clear (0) or almost clear (1) (on a 5 point scale) and a reduction from baseline of ≥2 points at Week 12. The baseline will be defined as the IGA score on Day 1 pre-dose.
		• Response based on the Eczema Area and Severity Index ≥75% improvement from baseline (EASI-75) response at Week 12. The baseline will be defined as the EASI score on Day 1 pre-dose.

CLINICAL STUDY REPORT SYNOPSIS

Table S1. Study Objectives and Endpoints

Secondary		
Efficacy	• To evaluate the effect of PF-04965842 on additional efficacy endpoints and patient-reported outcomes over time in subjects aged 12 years and older with moderate to severe atopic dermatitis.	**Key Secondary Efficacy Endpoints:** • Response based on at least 4 points improvement in the severity of pruritus numerical rating scale (NRS) from baseline at Weeks 2, 4, and 12.
		• Change from Baseline in Pruritus and Symptoms Assessment for Atopic Dermatitis (PSAAD) total score at Week 12.
		Secondary Efficacy Endpoints: • Response based on at least 4 points improvement in the severity of pruritus numerical rating scale, (described interchangeably as Peak Pruritus Numerical Rating Scale [PP-NRS]) from baseline at all scheduled time points other than Weeks 2, 4 and 12.
		• Time from baseline to achieve at least 4 points improvement in the severity of pruritus NRS scale.
		• Response based on the EASI-75 at all scheduled time points except Week 12.
		• Response based on the IGA of clear (0) or almost clear (1) and ≥2 point reduction from baseline at all scheduled time points except Week 12.
		Other Efficacy Endpoints: • Response based on a ≥50% and ≥90% improvement in the EASI total score (EASI-50 and EASI-90) at all scheduled time points.
		• Change from baseline in the percentage Body Surface Area (BSA) affected at all scheduled time points.
		• Response based on a ≥50% and ≥75% improvement in Scoring Atopic Dermatitis (SCORAD) (SCORAD50, SCORAD75) from baseline at all scheduled time points.
		• Change from baseline at all scheduled time points in SCORAD subjective assessments of itch and sleep loss.[d]
Patient-Reported Outcomes		• Change from baseline at Week 12 in Dermatology Life Quality Index (DLQI)

CLINICAL STUDY REPORT SYNOPSIS

Table S1. Study Objectives and Endpoints

		or Children's DLQI (CDLQI) and at all other scheduled time points.
		• Change from baseline at Week 12 in Hospital Anxiety and Depression Scale (HADS) and at all other scheduled time points.
		• Change from baseline at Week 12 in Patient-Oriented Eczema Measure (POEM) and at all other scheduled time points.
		• Change from baseline of Patient Global Assessment (PtGA) at Week 12 and at all other scheduled time points.
		• Change from baseline of EuroQol Quality of Life 5-Dimension 5-Level Scale (EQ-5D-5L) or EuroQol Quality of Life 5-Dimension Youth Scale (EQ-5D-Y) at Week 12 and at all other scheduled time points.
		• Change from baseline of Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F) or Pediatric FACIT-F (Peds-FACIT-F) at Week 12 and at all other scheduled time points.
		• Change from baseline of Short Form-36v2 (SF-36v2), acute at Week 12 and at all other scheduled time points.
		• Change from baseline of Work Productivity and Activity Impairment Questionnaire: Atopic Dermatitis (WPAI:AD) at Week 12 and at all other scheduled time points.
		• Response based on at least 4 points improvement in the frequency of pruritus numerical rating scale (NRS) from baseline at all scheduled time points.[a]
		• Time from baseline to achieve at least 4 points improvement in the frequency of pruritus NRS scale.[a]
Safety	• To evaluate the safety and tolerability of PF-04965842 in subjects aged 12 years and older with moderate to severe	• Incidence of treatment emergent adverse events.
		• Incidence of serious adverse events (SAEs) and adverse events (AEs) leading to discontinuation.

CLINICAL STUDY REPORT SYNOPSIS

Table S1. Study Objectives and Endpoints

	atopic dermatitis following 12 weeks of treatment.	• The incidence of clinical abnormalities and change from baseline in clinical laboratory values, electrocardiogram (ECG) measurements, and vital signs.
Pharmacokinetics		
Pharmacokinetics (PK)	• To evaluate the PK of PF-04965842 in subjects aged 12 years and older with moderate to severe atopic dermatitis following 12 weeks of treatment.	• Population PK characterization in subjects aged 12 years and older with moderate to severe atopic dermatitis.[b]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

a. The frequency of pruritus numerical rating scale (NRS) is listed as an endpoint in the protocol but was not analyzed. This was because, after its selection as an endpoint in the study protocol, evidence was uncovered that suggested the instrument did not have sufficient content validity. Specifically, the response scale for the frequency of pruritus NRS was found not to be ideally representative of the range of frequency levels from the perspective of the target population based on qualitative research. Data generated from this measure would not be easily interpretable. Therefore, the endpoint was removed from the statistical analysis plan (SAP).

b. A population PK model will be developed for estimating PK parameters. Additional details of the methodology and the results will be reported separately and not included in this CSR.

[redacted]

d. Change from baseline at all scheduled time points in SCORAD subjective assessments of itch was not analyzed because assessments of itch were measured more effectively in other endpoints of itch, such as the severity of itch (pruritus) due to AD assessed using the Pruritus Numerical Rating Scale.

CLINICAL STUDY REPORT SYNOPSIS

METHODS

Study Design: This was a randomized, double-blind, placebo-controlled, parallel-group, Phase 3 study to evaluate the efficacy and safety of PF-04965842 (henceforth referred to as abrocitinib) monotherapy in subjects aged 12 years and older with moderate-to-severe AD and a body weight ≥40 kg. The treatment duration was 12 weeks. Subjects were screened within 28 days prior to the first dose of investigational product to confirm that they met the subject selection criteria for the study. A total of 391 subjects were enrolled from 106 sites located globally.

Subjects who continued to meet eligibility criteria at baseline underwent Day 1/ baseline assessments and were randomized in a 2:2:1 ratio to receive abrocitinib 200 mg once daily (QD), abrocitinib 100 mg QD, or placebo. Randomization was stratified by baseline disease severity (moderate [IGA = 3] vs. severe [IGA = 4] AD), and age <18 and ≥18 years of age.

Diagnosis and Main Criteria for Inclusion: The study population consisted of subjects aged ≥12 years who met all the following AD criteria:

- Clinical diagnosis of chronic AD for at least 1 year prior to Day 1 and had confirmed AD (Hanifin and Rajka criteria of AD) at the screening and baseline visits;
- Documented recent history (within 6 months before the screening visit) of inadequate response to treatment with topical medications for at least 4 weeks, or for whom topical treatments are otherwise medically inadvisable (eg, because of important side effects or safety risks), or who have required systemic therapies for control of their disease;
- Moderate to severe AD (affected BSA ≥10%, IGA ≥3, EASI ≥16, and pruritus numeric rating scale [NRS] ≥4 on the day of the baseline visit).

Study Treatment: Subjects were dispensed 2 bottles at each dispensing visit and were to take one tablet from each bottle, once daily, preferably in the morning, at approximately the same time of day for 12 weeks (Table S2).

CLINICAL STUDY REPORT SYNOPSIS

Table S2. Investigational Product Description

Investigational Product Description	Vendor Lot Number	Pfizer Lot Number	Strength/Potency	Dosage Form
PF-04965842 100 mg round white film coated tablet (DC)	[redacted]	[redacted]	100 mg	Tablet
PF-04965842 100 mg round white film coated tablet (DC)	[redacted]	[redacted]	100 mg	Tablet
PF-04965842 100 mg round white film coated tablet (DC)	[redacted]	[redacted]	100 mg	Tablet
Placebo for PF-04965842 round white film coated tablet (9 mm)	[redacted]	[redacted]	0 mg	Tablet
Placebo for PF-04965842 round white film coated tablet (9 mm)	[redacted]	[redacted]	0 mg	Tablet

Efficacy Evaluations:

The co-primary endpoints were:

- Response based on the IGA score of clear (0) or almost clear (1); and a reduction from baseline of ≥2 points at Week 12.
- Response based on the EASI ≥75% improvement from baseline (EASI-75) at Week 12.

The key secondary efficacy endpoints analyzed were:

- Response based on ≥4 points improvement from baseline in the pruritus NRS (NRS4) for severity at Weeks 2, 4, and 12;
- Change from baseline in PSAAD at Week 12.

The secondary efficacy endpoints analyzed were the following:

- Response based on NRS4 for severity at all time points other than Weeks 2, 4, and 12.

CLINICAL STUDY REPORT SYNOPSIS

- Time to achieve a ≥4-point improvement from baseline in the numerical rating scale (NRS) for severity of pruritus.
- Response based on the EASI-75 at all scheduled time points except Week 12.
- Response based on the IGA for clear (0) or almost clear (1); and ≥2 points reduction from baseline at all scheduled time points except Week 12.

The other secondary efficacy endpoints analyzed were the following:

- Response based on the IGA for clear (0) at all scheduled time points.
- Response based on a ≥50%, ≥90%, and =100% improvement in the EASI total score (EASI-50, EASI-90, and EASI-100) at all scheduled time points.
- Percent change from baseline at all scheduled time points in the EASI total score.
- Change from baseline in %BSA (from EASI) affected at all scheduled time points.
- Proportion of patients with %BSA (from EASI) <5% at all scheduled time points.
- Response based on a ≥50% and ≥75% improvement in SCORAD (SCORAD50, SCORAD75) from baseline at all scheduled time points.
- Percent change from baseline at all scheduled time points in the SCORAD total score.
- Change from baseline at all scheduled time points in SCORAD subjective visual analog scale (VAS) assessments of sleep loss.

The patient-reported outcome (PRO) secondary endpoints analyzed were the following:

- Change from baseline in DLQI/CDLQI at all scheduled time points.
- Change from baseline in each component (anxiety and depression) of the HADS score at all scheduled time points.
- Change from baseline in POEM at all scheduled time points.
- Change from baseline in PtGA at all scheduled time points.
- Change from baseline in EQ-5D-5L/EQ-5D-Y at all scheduled time points.
- Change from baseline in FACIT-F/Peds-FACIT-F at Week 12.

CLINICAL STUDY REPORT SYNOPSIS

- Change from baseline in the physical and mental component scores of the SF-36v2, Acute at Week 12.
- Change from baseline of Work Productivity and Activity Impairment Questionnaire: Atopic Dermatitis (WPAI:AD) at Week 12.
- Response based on achieving ≥4-point improvement from baseline in the DLQI score at all scheduled time points (among subjects with a score ≥4 at baseline).
- Response based on achieving a DLQI score <2 at all scheduled time points (among adult subjects with a score ≥2 at baseline).
- Response based on PtGA score of clear (0) or almost clear (1) with a reduction from baseline of ≥2 points at all scheduled time points (among subjects with a score ≥2 at baseline).
- Response based on achieving endpoint NRS severity itch score ≤ 2 for patients with baseline score >2.
- Response based on achieving ≥1 (clinically important response or CIR) point improvement from baseline in the total PSAAD score at all scheduled time points.
- Change from baseline in PSAAD score at all scheduled time points except Week 12.

[redacted]

Pharmacokinetic Evaluations: The PK endpoint was population PK characterization in subjects aged 12 years and older with moderate to severe atopic dermatitis. Blood for PK analysis was collected at the study site at 2.0 hours (±30 minutes) pre-dose at Week 8 and at 1.0 hour (±15 minutes) and 2.0 hours (±30 minutes) post-dose at Week 12.

Safety Evaluations: Safety was assessed by the spontaneous reporting of AEs, physical examinations, vital signs, and clinical laboratory results in all subjects who received at least 1 dose of the investigational product. The safety endpoints included incidence of treatment-emergent AEs (TEAEs), incidence of SAEs and AEs leading to discontinuation, and incidence of clinical abnormalities and change from baseline in clinical laboratory values, ECG measurements, and vital signs.

CLINICAL STUDY REPORT SYNOPSIS

Statistical Methods: Abrocitinib 200 mg QD was declared superior to placebo if the null hypothesis of no difference between abrocitinib 200 mg QD versus placebo for both co-primary endpoints was rejected at the 5% significance level. Similarly, abrocitinib 100 mg QD was declared superior to placebo if the null hypothesis of no difference between abrocitinib 100 mg QD versus placebo for both co-primary endpoints was rejected at the significance level specified below.

A sequential Bonferroni-based iterative multiple testing procedure to strongly control the familywise Type 1 error at 5% was used for testing each of the two abrocitinib doses (200 mg QD and 100 mg QD) versus placebo on the primary and key secondary endpoints. The procedure belonged to a class of consonant multiple test procedures, which are a subclass of the closed test procedures.

For co-primary and secondary endpoints, binary data at each scheduled visit were analyzed. The test of hypothesis between the abrocitinib treatment groups versus the placebo group were conducted by the Cochran-Mantel-Haenszel (CMH) statistic adjusting for the effect of randomization strata; p-values from the CMH statistic were used to establish the superiority of each dose of abrocitinib to placebo in binary responses. The proportion of responders in the abrocitinib treatment groups versus the placebo group were summarized by the difference and its 95% confidence interval (CI) obtained by normal approximation. The difference in proportions was calculated within each randomization stratum. The final estimate of the difference in proportions was a weighted average of these stratum-specific estimates using CMH weights. Estimates of the difference in proportions along with the two-sided 95% CI were also provided for the abrocitinib 200 mg QD group versus the abrocitinib 100 mg QD group. No hypotheses were tested. In this analysis, subjects who permanently discontinued study for any reason were defined as "non-responders" at all subsequent visits.

For secondary endpoints, the proportion of responders in each treatment group along with the differences in proportion of responders among each pair of treatment groups were obtained using the methods described above. To account for the responses which were missing because the data could not be assessed at the scheduled visits of Weeks 2, 4, 8, and 12 in several subjects, a hybrid approach was used. First, an imputation model based on a generalized linear mixed model (GLMM) was fit to the observed data with treatment, visit, and treatment-by-visit interaction as fixed factors and a subject-specific normally-distributed random effect. Second, any responses which were missing due to the subject discontinuing permanently or due to a rescue medication were defined as "non-response". Third, any other response which remained missing at any intermittent visits was multiply imputed under the missing at random (MAR) assumption. The multiple imputation methodology was used with the missing not at random (MNAR) weights taken as zero (ie, assuming MAR).

The change from baseline in the PSAAD score was analyzed as longitudinal continuous data. A mixed-effect, repeated measures (MMRM) model was used. The fixed effects of treatment, visit, treatment-by-visit interaction, and randomization stratification factors were included. Visit was modeled as a categorical covariate. An unstructured covariance matrix

CLINICAL STUDY REPORT SYNOPSIS

was assumed for the model errors. A compound symmetry covariance matrix was used if the model with unstructured covariance did not converge.

When modeling the change from baseline values, the variable for visit started with the first postbaseline visit, and the actual baseline value was included as a covariate. At each visit, estimates of least squares mean (LSM) values and the LSM differences between the abrocitinib treated groups and the placebo group were derived from the model. The corresponding p-values, standard errors, and 95% CIs were also derived from the model. Estimates of the difference in LSMs along with the two-sided 95% CI were also provided for the abrocitinib 200 mg QD group versus the abrocitinib 100 mg QD group. No hypotheses were tested. The MMRM yielded valid inferences in the presence of missing data mechanism which was assumed to be MAR.

RESULTS

Subject Disposition and Demography: A total of 554 subjects were screened and 391 subjects were randomized. Higher proportions of subjects in the abrocitinib treatment groups completed the study compared with the placebo group (Table S3). The proportion of subjects who discontinued was higher in the placebo group compared with either abrocitinib treatment group; this was primarily accounted for by a higher proportion of subjects who discontinued due to withdrawal by subject, AEs, and lack of efficacy.

Table S3. Disposition Events Summary

Number (%) of Subjects	Placebo (N=78) n (%)	PF-04965842 100mg QD (N=158) n (%)	PF-04965842 200mg QD (N=155) n (%)
Discontinued	26 (33.3)	21 (13.3)	14 (9.0)
Adverse Event	8 (10.3)	5 (3.2)	5 (3.2)
Death	0	1 (0.6)	0
Lack of Efficacy	7 (9.0)	5 (3.2)	4 (2.6)
Lost to Follow-Up	1 (1.3)	1 (0.6)	1 (0.6)
Protocol Deviation	1 (1.3)	1 (0.6)	1 (0.6)
Withdrawal By Subject	9 (11.5)	6 (3.8)	1 (0.6)
Other	0	2 (1.3)	2 (1.3)
Completed	52 (66.7)	137 (86.7)	141 (91.0)

Completed means completed 12 weeks treatment.
[redacted] SDTM Creation: 12SEP2019 (22:25) Source Data: adds Output File: ./nda1 cdisc/B7451013/adds s001 Date of Generation: 18OCT2019 (00:14)
Table 14.1.1.2 is for Pfizer internal use.

CLINICAL STUDY REPORT SYNOPSIS

The analyses of the efficacy and PRO endpoints were performed using the Full Analysis Set (FAS) population, which was defined as all 391 randomized subjects who received at least 1 dose of study medication. Additional supportive analyses of the primary and key secondary efficacy endpoints were performed for the per protocol analysis set (PPAS), which excluded 25, 30, and 26 subjects in the abrocitinib 200 mg QD, abrocitinib 100 mg QD, and placebo groups, respectively, resulting in an overall 79.2% of the total FAS population. The analysis of AEs and laboratory data was performed for the Safety Analysis Set, which comprised all subjects who received at least 1 dose of study medication and was identical to the FAS population.

Baseline demographic and disease characteristics were balanced across treatment groups. The median age of subjects randomized across treatment groups was 31.0 years, with approximately 10% of subjects who were adolescents aged 12 to <18 years at the time of screening. The majority of subjects across treatment groups were white (59.3%); 33.0% of all subjects were of Asian descent and 5.4% were Black or African American. The median disease duration (Q1, Q3) was approximately 19.6 years (9.2, 29.2). Randomized subjects were representative of the moderate to severe AD population, with all subjects with moderate (67.8%) or severe (32.3%) IGA, median EASI score of 25.2, median PP-NRS of 7.0, and baseline PRO values representative of a high burden of disease and impact on quality of life (QoL).

Efficacy Results:

Co-Primary Endpoints

The study met both co-primary endpoints of IGA response at Week 12 and EASI-75 response at Week 12 demonstrating that both abrocitinib 200 mg QD and 100 mg QD treatment groups were superior to the placebo group (Table S4 and Table S5). Additional supportive sensitivity analysis results based on the PPAS and a tipping point analysis (where all missing responses were multiply imputed) were consistent with the FAS primary analysis.

CLINICAL STUDY REPORT SYNOPSIS

Table S4. Proportion of Subjects Achieving Investigator's Global Assessment (IGA) Response of 'Clear' or 'Almost Clear' and >=2 Points Improvement from Baseline at Week 12 - CMH (FAS, NRI)

	Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD
N	77	155	155
n (%)	7 (9.1)	44 (28.4)	59 (38.1)
95% CI	(2.7, 15.5)	(21.3, 35.5)	(30.4, 45.7)
Active - Placebo [1]			
Estimate (%)		19.3	28.7
95% CI		(9.6, 29.0)	(18.6, 38.8)
Two-sided P-value [2]		0.0008	<.0001
200 mg QD - 100 mg QD [1]			
Estimate (%)			9.7
95% CI			(-0.7, 20.0)

Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the last measurement prior to first dosing (Day 1).
If a subject withdrew from the study, then this subject was counted as non-responder after withdrawal.
CI = confidence interval; CMH = Cochran-Mantel-Haenszel; N = number of subjects in the analysis set with NRI at the specified visit; n (%) = number of subjects with response (percentage based on N); NRI = non-responder imputation.
[1] The estimate and confidence interval (CI) for difference were calculated based on the weighted average of difference for each randomization stratum using the
normal approximation of binomial proportions. The confidence interval for the response rate was based on normal approximation (or the Clopper-Pearson exact method when there were no or if all were responders).
[2] P-value was calculated using the Cochran-Mantel-Haenszel (CMH) method adjusted by randomization strata (baseline disease severity and age category).
[redacted] SDTM Creation: 12SEP2019 (22:27) Source Data: adad Output
File: ./nda1 cdisc/B7451013/adad s102 Date of Generation: 18OCT2019 (05:44)
Table 14.2.1.1.1.1 is for Pfizer internal use.

CLINICAL STUDY REPORT SYNOPSIS

Table S5. Proportion of Subjects Achieving Eczema Area and Severity Index (EASI) Response >= 75% Improvement from Baseline at Week 12 - CMH (FAS, NRI)

	Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD
N	77	155	154
n (%)	8 (10.4)	69 (44.5)	94 (61.0)
95% CI	(3.6, 17.2)	(36.7, 52.3)	(53.3, 68.7)
Active - Placebo [1]			
Estimate (%)		33.9	50.5
95% CI		(23.3, 44.4)	(40.0, 60.9)
Two-sided P-value [2]		<.0001	<.0001
200 mg QD - 100 mg QD [1]			
Estimate (%)			16.5
95% CI			(5.6, 27.4)

Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the last measurement prior to first dosing (Day 1).
If a subject withdrew from the study, then this subject was counted as non-responder after withdrawal.
CI = confidence interval; CMH = Cochran-Mantel-Haenszel; N = number of subjects in the analysis set with NRI at the specified visit; n (%) = number of subjects with response (percentage based on N); NRI = non-responder imputation.
[1] The estimate and confidence interval (CI) for difference were calculated based on the weighted average of difference for each randomization stratum using the
normal approximation of binomial proportions. The confidence interval for the response rate was based on normal approximation (or the Clopper-Pearson exact method when there were no or if all were responders).
[2] P-value was calculated using the Cochran-Mantel-Haenszel (CMH) method adjusted by randomization strata (baseline disease severity and age category).
[redacted] SDTM Creation: 12SEP2019 (22:46) Source Data: adea Output File: ./nda1 cdisc/B7451013/adea s104 Date of Generation: 18OCT2019 (06:00)
Table 14.2.1.1.3.1 is for Pfizer internal use.

Key Secondary Endpoints

Both key secondary endpoints were met. At Weeks 2, 4, and 12, both abrocitinib treatment groups had statistically significantly greater proportions of subjects achieving ≥4 points improvement from baseline in severity of pruritus NRS (PP-NRS4) responders compared with the placebo group (Table S6). Separation from placebo in PP-NRS4 response began as early as Week 2 in the abrocitinib 200 mg QD and 100 mg QD treatment groups.
At Week 12, the LSM of change from baseline in PSAAD scores showed statistically significant decreases from baseline for both abrocitinib treatment groups compared with the placebo group (Table S7).

CLINICAL STUDY REPORT SYNOPSIS

Table S6. Proportion of Subjects with >=4 Points at Baseline and Achieving >=4 Points Improvement from Baseline in Numeric Rating Scale for Severity of Pruritus - CMH (FAS, NRI after Dropout + MI for Intermittent Missing)

		Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD
Week 2	N	76	156	153
	Estimated Response Rate (%)	3.9	23.1	35.3
	95% CI	(0.0, 8.3)	(16.5, 29.7)	(27.7, 42.9)
	Active - Placebo			
	Estimate (%)		19.2	31.2
	95% CI		(11.0, 27.4)	(22.3, 40.2)
	Two-sided P-value		0.0002	<.0001
	200 mg QD - 100 mg QD			
	Estimate (%)			12.1
	95% CI			(2.2, 22.1)
Week 4	N	76	156	153
	Estimated Response Rate (%)	4.0	33.4	52.8
	95% CI	(0.0, 8.4)	(25.8, 41.0)	(44.7, 60.8)
	Active - Placebo			
	Estimate (%)		29.5	48.8
	95% CI		(20.5, 38.4)	(39.5, 58.2)
	Two-sided P-value		<.0001	<.0001
	200 mg QD - 100 mg QD			
	Estimate (%)			19.4
	95% CI			(8.4, 30.4)
Week 8	N	76	156	153
	Estimated Response Rate (%)	12.0	40.4	54.4
	95% CI	(4.6, 19.4)	(32.6, 48.2)	(46.4, 62.4)
	Active - Placebo			
	Estimate (%)		28.5	42.4
	95% CI		(17.8, 39.3)	(31.4, 53.4)
	Two-sided P-value		<.0001	<.0001
	200 mg QD - 100 mg QD			
	Estimate (%)			14.0
	95% CI			(2.9, 25.1)
Week 12	N	76	156	153
	Estimated Response Rate (%)	11.5	45.2	55.3
	95% CI	(4.1, 19.0)	(37.1, 53.3)	(47.2, 63.5)
	Active - Placebo			
	Estimate (%)		33.7	43.9
	95% CI		(22.8, 44.7)	(32.9, 55.0)
	Two-sided P-value		<.0001	<.0001

CLINICAL STUDY REPORT SYNOPSIS

Table S6. Proportion of Subjects with >=4 Points at Baseline and Achieving >=4 Points Improvement from Baseline in Numeric Rating Scale for Severity of Pruritus - CMH (FAS, NRI after Dropout + MI for Intermittent Missing)

	Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD
200 mg QD - 100 mg QD			
Estimate (%)			10.2
95% CI			(-1.1, 21.5)

Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the measurement collected on or prior to Day 1.
If a subject withdrew from the study, then this subject was counted as non-responder after withdrawal. Any other intermittent missing value was handled using multiple imputation.
CI = confidence interval; CMH = Cochran-Mantel-Haenszel; GLMM = generalized linear mixed model; MI = multiple imputation; N = number of subjects in the analysis set with NRI + MI at the specified visit with baseline >= 4; NRI = non-responder imputation.
At Week 2, no multiple imputations were performed because there were <= 5 subjects with missing responses in all treatment groups.
[redacted] SDTM Creation: 12SEP2019 (22:28) Source Data: adnr mif Output File: ./nda1_cdisc/B7451013/adnr_s201 Date of Generation: 18OCT2019 (04:00)
Table 14.2.3.3.1 is for Pfizer internal use.

Table S7. Least Squares Mean of Change from Baseline in Pruritus and Symptoms Assessment for Atopic Dermatitis (PSAAD) at Week 12 - MMRM (FAS, OD)

	Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD
N	77	156	155
LSM	-0.8	-2.4	-3.0
95% CI	(-1.3, -0.3)	(-2.8, -2.1)	(-3.3, -2.7)
Active - Placebo			
LSM		-1.7	-2.2
95% CI		(-2.3, -1.1)	(-2.8, -1.6)
P-value		<.0001	<.0001
200 mg QD - 100 mg QD			
LSM			-0.6
95% CI			(-1.0, -0.1)

CLINICAL STUDY REPORT SYNOPSIS

Table S7. Least Squares Mean of Change from Baseline in Pruritus and Symptoms Assessment for Atopic Dermatitis (PSAAD) at Week 12 - MMRM (FAS, OD)

	Placebo	PF-04965842 100mg QD	PF-04965842 200mg QD

Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the average of all values recorded between Day -6 and Day 1.
Weekly data were average values of daily observations over 7 days.
CI = confidence interval; LSM = least squares mean; N = number of subjects included in the analysis model; OD = observed data.
Mixed Model Repeated Measure (MMRM) contained fixed factors of treatment, visit, treatment by visit interaction, randomization strata (baseline disease severity and age category), baseline value and an unstructured covariance matrix.
[redacted] SDTM Creation: 12SEP2019 (22:43) Source Data: adpu Output
File: ./nda1 cdisc/B7451013/adpu s103 Date of Generation: 18OCT2019 (10:27)
Table 14.2.4.2.3 is for Pfizer internal use.

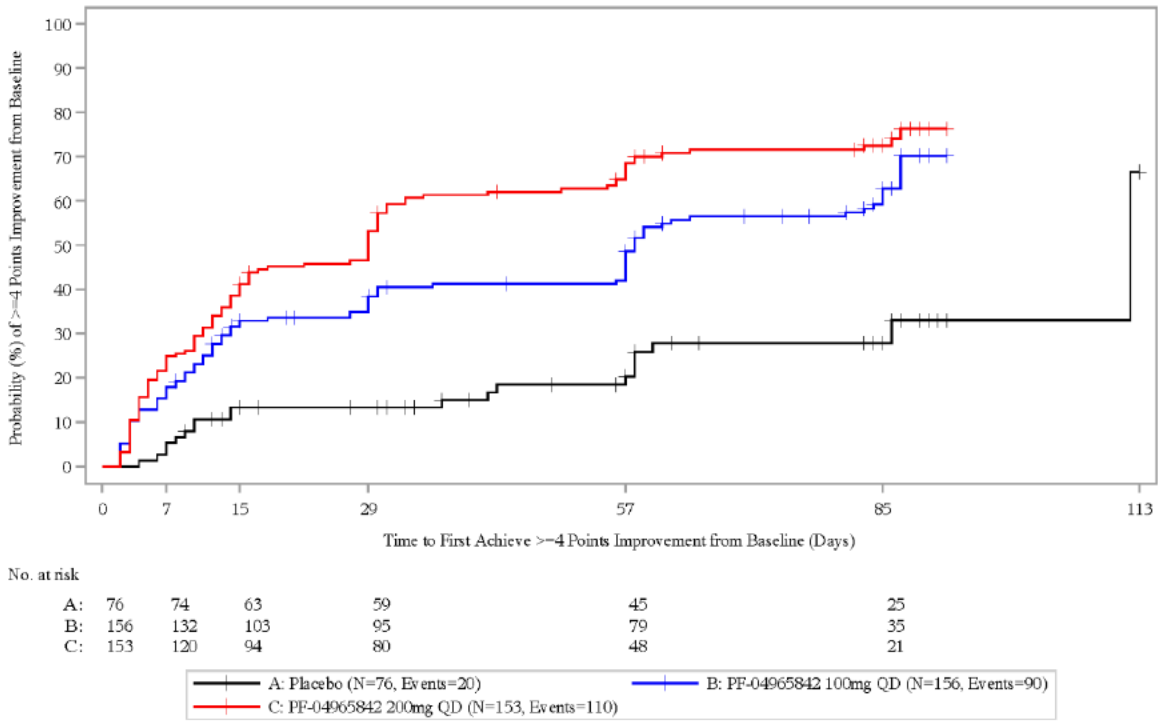
Secondary Efficacy Endpoints

The secondary efficacy endpoints showed improved efficacy in both abrocitinib treatment groups compared with the placebo group.

- At Week 8, both abrocitinib treatment groups had greater proportions of PP-NRS4 responders compared with the placebo group (Table S6).
- The Kaplan-Meier analysis to estimate the time to first PP-NRS4 showed shorter median times in both abrocitinib treatment groups compared with the placebo group (Figure S1).

CLINICAL STUDY REPORT SYNOPSIS

Figure S1. Kaplan-Meier Plot of Time to First Achieve ≥4 Points Improvement from Baseline in Numeric Rating Scale for Severity of Pruritus (FAS)

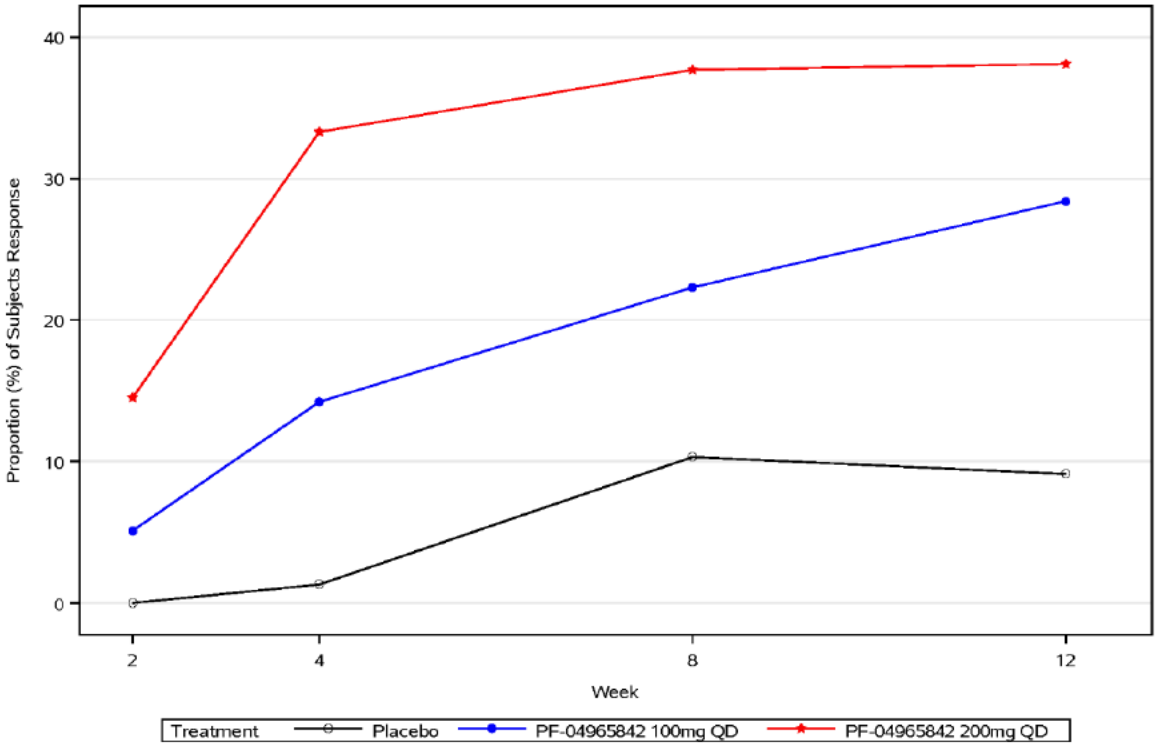


Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
N = number of subjects from the FAS population with a baseline numeric rating scale score for severity of pruritus >=4.
Baseline was defined as the measurement collected on or prior to Day 1.
Follow-up time for each subject is from baseline until the last observation.
[redacted] SDTM Creation: 12SEP2019 (22:28) Source Data: adtte Output File: ./nda1_cdisc/B7451013/adnr_f301 Date of Generation: 17OCT2019 (23:42)

- At Weeks 2, 4, and 8, both abrocitinib treatment groups showed higher placebo-corrected proportions of subjects achieved IGA responses compared with the placebo group (Figure S2).

CLINICAL STUDY REPORT SYNOPSIS

Figure S2. Plot of Proportion of Subjects Achieving Investigator's Global Assessment (IGA) Response of 'Clear' or 'Almost Clear' and ≥2 Points Improvement from Baseline (FAS, NRI)

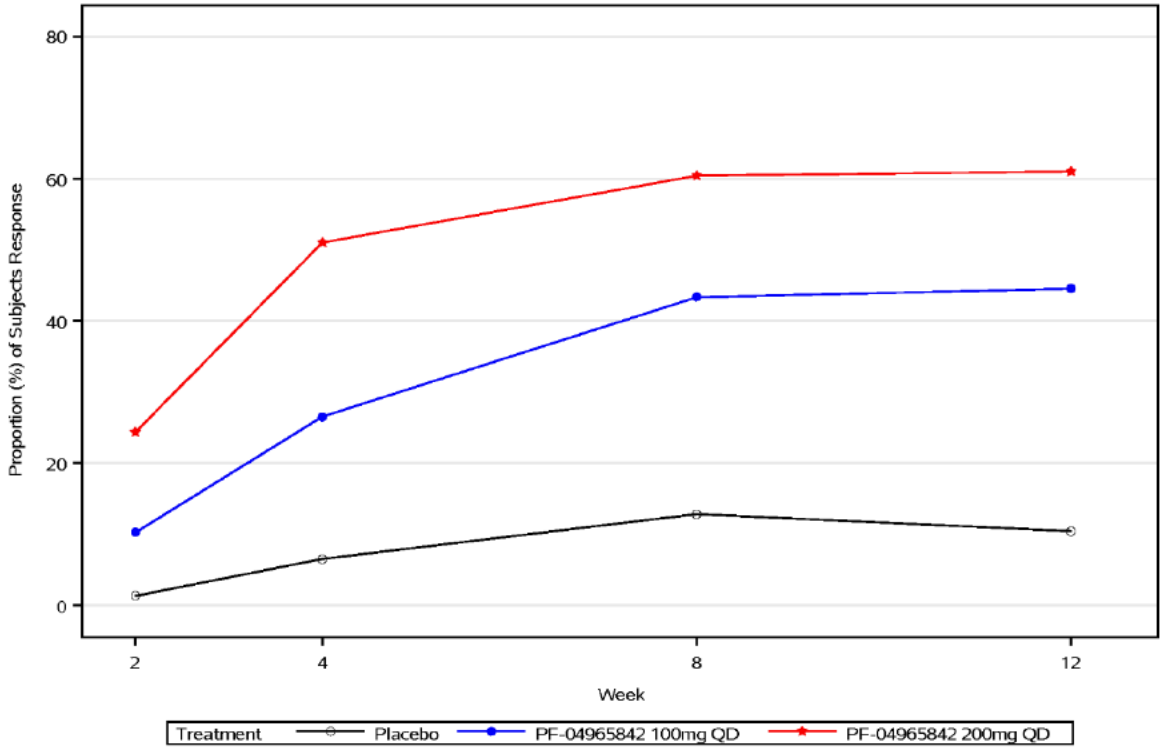


Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the last measurement prior to first dosing (Day 1).
If a subject withdrew from the study, then this subject was counted as non-responder after withdrawal.
NRI = non-responder imputation.
[redacted] SDTM Creation: 12SEP2019 (22:27) Source Data: adad Output File: ./nda1_cdisc/B7451013/adad_f101 Date of Generation: 18OCT2019 (22:08)

- At Weeks 2, 4, and 8, both abrocitinib treatment groups showed higher placebo-corrected proportions of subjects achieved EASI-75 responses compared with the placebo (Figure S3).

CLINICAL STUDY REPORT SYNOPSIS

Figure S3. Plot of Proportion of Subjects Achieving Eczema Area and Severity Index (EASI) Response ≥75% Improvement from Baseline (FAS, NRI)



Full Analysis Set (FAS) was defined as all randomized subjects who received at least one dose of study medication.
Baseline was defined as the last measurement prior to first dosing (Day 1).
If a subject withdrew from the study, then this subject was counted as non-responder after withdrawal.
NRI = non-responder imputation.
[redacted] SDTM Creation: 12SEP2019 (22:46) Source Data: adea Output File: ./nda1_cdisc/B7451013/adea_f102 Date of Generation: 18OCT2019 (22:14)

Other Secondary Endpoints

EASI-50/EASI-90: Both abrocitinib treatment groups demonstrated higher EASI-50 responses compared with the placebo group as early as Week 2 that were maintained up to Week 12. Beginning at Week 2 in the abrocitinib 200 mg QD group and Week 4 in the abrocitinib 100 mg QD group, higher EASI-90 responses compared with the placebo group were observed that were maintained up to Week 12.

Percentage BSA: Both abrocitinib treatment groups demonstrated decreased absolute and percent change from baseline in %BSA compared with the placebo group as early as Week 2 that were maintained up to Week 12.

SCORAD50/SCORAD75: Both abrocitinib treatment groups demonstrated higher SCORAD50 responses compared with the placebo group as early as Week 2 that were maintained up to Week 12. Both abrocitinib treatment groups demonstrated higher SCORAD75 responses compared with the placebo group beginning at Week 2 for the abrocitinib 200 mg QD group and at Week 4 for the abrocitinib 100 mg QD group.

CLINICAL STUDY REPORT SYNOPSIS

SCORAD Sleep Loss: Both abrocitinib treatment groups demonstrated decreases in SCORAD VAS of sleep loss absolute and percent change from baseline compared with the placebo group beginning at Week 2 that were maintained up to Week 12, except for Weeks 2 and 12 in the abrocitinib 100 mg QD group.

Patient-Reported Outcomes

DLQI/CDLQI: For DLQI (adult subjects), both abrocitinib treatment groups showed decreased DLQI scores compared with the placebo group that occurred as early as Week 2 and were maintained up to Week 12. For CDLQI (adolescent subjects), the sample sizes were low (15 subjects in the abrocitinib 200 mg QD group, 16 subjects in the abrocitinib 100 mg QD group, and 8 subjects in the placebo group) and showed decreased CDLQI scores compared with the placebo group beginning as early as Week 2 in the abrocitinib 200 mg QD group and at no time point in the abrocitinib 100 mg QD group.

HADS: The LSM of change from baseline in HADS anxiety sub-scores was improved (lower) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 and maintained at all subsequent time points, except at Week 8 in the abrocitinib 100 mg QD group. The LSM of change from baseline in HADS depression sub-scores was improved (lower) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 and maintained at all subsequent time points.

POEM: The LSM of change from baseline in POEM scores was improved (lower) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 that were maintained up to Week 12.

PtGA: The LSM of change from baseline in PtGA scores was improved (lower) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 that were maintained up to Week 12.

EQ-5D-5L/EQ-5D-Y: For the EQ-5D-5L index value, the LSM of change from baseline was improved (higher) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 that were maintained at all scheduled time points. For the EQ-5D-5L EQ VAS score, the LSM of change from baseline was improved (higher) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 that were maintained at all scheduled time points, except for Week 8 in the abrocitinib 100 mg QD group.

FACIT-F/ Peds-FACIT-F: The LSM of change from baseline for FACIT-F in adult subjects was improved (higher scores) for both abrocitinib treatment groups compared with the placebo group at Week 12. The LSM of change from baseline for Peds-FACIT-F in adolescent subjects was not different for both abrocitinib treatment groups compared with the placebo group at Week 12; however, the number of adolescent subjects was low in all treatment groups.

CLINICAL STUDY REPORT SYNOPSIS

SF-36v2: The LSM of change from baseline in the physical component summary score in SF-36v2 was improved (higher) in both abrocitinib treatment groups compared with the placebo group at Week 12. The LSM of change from baseline in the mental component summary score in SF-36v2 was improved (higher) in the abrocitinib 200 mg QD group only compared with the placebo group at Week 12.

WPAI:AD: The LSM of change from baseline in the WPAI:AD (including percent impairment while working, percent overall work impairment, and percent activity impairment) was improved (lower) in both abrocitinib treatment groups compared with the placebo group at Week 12. The LSM of change from baseline in the WPAI:AD (percent work time missed) was not different in both abrocitinib treatment groups compared with the placebo group at Week 12.

[redacted]

Pharmacokinetic Results: For both abrocitinib 200 mg QD and 100 mg QD groups, mean plasma abrocitinib concentrations appeared to be highest at the 1-hour post-dose PK sampling time in this study. The mean plasma abrocitinib concentrations observed prior to dosing (Week 8), and at 1-hour and 2-hours post-dose (Week 12) increased in a proportion approximately similar to the 2-fold increase in the dose between the abrocitinib 200 mg QD group and abrocitinib 100 mg QD group.

Safety Results:

- A higher incidence of subjects with all-causality TEAEs occurred in the abrocitinib treatment groups compared with the placebo group (Table S8).
    - The majority of reported all-causality TEAEs were mild or moderate; the incidence of severe TEAEs was similar between the abrocitinib treatment groups and the placebo group.
    - Approximately one-third to half of all-causality TEAEs were considered treatment-related by the investigator.
- A similar proportion of subjects had serious TEAEs across all treatment groups (1.3% to 3.2%); one death (in a subject in the abrocitinib 100 mg QD group) was reported in the study.

CLINICAL STUDY REPORT SYNOPSIS

- The incidence of subjects who permanently discontinued from the study due to TEAEs was lower in the abrocitinib treatment groups compared with the placebo group.

- The incidence of subjects with temporary discontinuations of study drug due to a TEAE was similar between the abrocitinib treatment groups and the placebo group.

Table S8. Treatment-Emergent Adverse Events (All Causalities)

Number (%) of Subjects	Placebo n (%)	PF-04965842 100mg QD n (%)	PF-04965842 200mg QD n (%)
Subjects evaluable for adverse events	78	158	155
Number of adverse events	70	204	245
Subjects with adverse events	42 (53.8)	99 (62.7)	102 (65.8)
Subjects with serious adverse events	1 (1.3)	5 (3.2)	2 (1.3)
Subjects with severe adverse events	5 (6.4)	7 (4.4)	6 (3.9)
Subjects discontinued from study due to adverse events [1]	10 (12.8)	6 (3.8)	5 (3.2)
Subjects discontinued study drug due to AE and continued Study [2]	0	2 (1.3)	0
Subjects with temporary discontinuation due to adverse events	2 (2.6)	8 (5.1)	5 (3.2)

Included data up to 28 days after last dose of study.
Except for the number of adverse events subjects were counted only once per treatment in each row.
Serious Adverse Events - according to the investigator's assessment
[1] Subjects who had an AE record that indicated that the AE caused the subject to be discontinued from the study
[2] Subjects who had an AE record that indicated that action taken with study treatment was drug withdrawn but AE did not cause the subject to be discontinued from study
MedDRA v22.0 coding dictionary applied.
[redacted] SDTM Creation: 12SEP2019 (22:27) Source Data: adae Output File: ./nda1 cdisc/B7451013/adae s010 Date of Generation: 18OCT2019 (00:24)
Table 14.3.1.2.1 is for Pfizer internal use.

CLINICAL STUDY REPORT SYNOPSIS

Conclusions:

EFFICACY

Co-Primary Endpoints:

- Both abrocitinib treatment groups (200 mg QD and 100 mg QD) met the co-primary endpoints of IGA response of clear (0) or almost clear (1) and ≥2-points improvement from baseline at Week 12 and EASI response ≥75% improvement from baseline at Week 12 compared with the placebo group.
    - Statistically significantly higher placebo-corrected proportions of subjects achieved IGA responses and EASI-75 responses for both abrocitinib treatment groups compared with the placebo group.
    - IGA responses and EASI-75 responses (placebo-corrected) in the abrocitinib 200 mg QD group were approximately 50% higher than the abrocitinib 100 mg QD group.
    - Sensitivity analyses of IGA and EASI-75 responses were consistent with the primary analyses.
    - The treatment effects for both IGA response and EASI-75 response were similar across subgroups.

Key Secondary Endpoints:

- Both abrocitinib treatment groups (200 mg QD and 100 mg QD) met the key secondary endpoints of PP-NRS4 response at Weeks 2, 4, and 12, and PSAAD change from baseline at Week 12.
    - Both abrocitinib treatment groups (200 mg QD and 100 mg QD) showed statistically significantly greater proportions of PP-NRS4 responders compared with the placebo group at Weeks 2, 4, and 12. Separation from placebo in PP-NRS4 response began as early as Week 2 in the abrocitinib 200 mg QD and 100 mg QD treatment groups that were maintained at all subsequent scheduled time points.
    - Both abrocitinib treatment groups (200 mg QD and 100 mg QD) showed statistically significant decreases from baseline in PSAAD scores compared with the placebo group at Week 12. As early as Week 1, the LSM of change from baseline for both abrocitinib treatment groups were decreased compared with the placebo group, and the treatment effect of both abrocitinib groups plateaued after Week 5.

CLINICAL STUDY REPORT SYNOPSIS

Other Secondary Endpoints:

- Additional secondary efficacy endpoints (IGA, EASI, PP-NRS, BSA, and SCORAD) demonstrated improved efficacy in both abrocitinib treatment groups compared with the placebo group.

IGA

- Both abrocitinib treatment groups showed higher IGA responses compared with placebo at Week 2, suggesting early onset of the treatment effect, and responses were maintained at Weeks 4 and 8.

EASI

- Both abrocitinib treatment groups demonstrated higher EASI-50 responses and EASI-90 responses compared with the placebo group.
- The higher EASI-50 responses and EASI-90 responses compared with the placebo group began as early as Week 2 and were maintained through Week 12.
- Both abrocitinib treatment groups had higher EASI-75 responses compared with placebo at Week 2, suggesting early onset of the treatment effect, and responses were maintained at Weeks 4 and 8.

PP-NRS

- The LSM of change from baseline in PP-NRS was lower in both abrocitinib groups compared with the placebo group as early as Day 2 (24 hours after the first dose), demonstrating a rapid onset of the treatment effect, which was maintained through Week 12.
- The median time to achieve first PP-NRS4 was faster in both abrocitinib treatment groups compared with the placebo group. In the abrocitinib 200 mg QD and 100 mg QD groups, 50% improved by 29 days and 58 days, respectively, compared with the placebo group.

BSA

- Both abrocitinib treatment groups demonstrated decreases in absolute and percent change from baseline in %BSA compared with the placebo group as early as Week 2 that were maintained up to Week 12.

CLINICAL STUDY REPORT SYNOPSIS

SCORAD

- Both abrocitinib treatment groups demonstrated higher SCORAD50 responses compared with the placebo group as early as Week 2 that were maintained up to Week 12.

- Both abrocitinib treatment groups demonstrated higher SCORAD75 responses compared with the placebo group beginning at Week 2 for the abrocitinib 200 mg QD group and at Week 4 for the abrocitinib 100 mg QD group.

- Both abrocitinib treatment groups demonstrated decreases in SCORAD total score absolute and percent change from baseline compared with the placebo group beginning at Week 2 that were maintained up to Week 12.

- Both abrocitinib treatment groups demonstrated decreases in SCORAD VAS of sleep loss absolute change from baseline and percent change from baseline compared with the placebo group beginning at Week 2 that were maintained up to Week 12, except for percent change from baseline at Weeks 2 and 12 in the abrocitinib 100 mg QD group.

- Across all secondary efficacy endpoints evaluated, the treatment effect of the abrocitinib 200 mg QD group was consistently higher than the abrocitinib 100 mg QD group.

Patient-Reported Outcomes (PROs):

- Favorable and improved differences across all PRO measures were observed for both abrocitinib treatment groups compared with the placebo group at Week 12. For all adult measures assessed (DLQI, HADS, POEM, PtGA, EQ-5D-5L, FACIT-F, SF-36v2, and WPAI:AD), treatment differences emerged as early as Week 2 and were maintained at subsequent time points up to Week 12.

    - There was an improvement in dermatology-related QoL in the DLQI assessments in both abrocitinib treatment groups compared with the placebo group.

    - An improvement in symptoms associated with depression and anxiety in the HADS sub-scores in subjects in both abrocitinib treatment groups compared with the placebo group was seen at all time points, except at Week 8 in the abrocitinib 100 mg QD group for the HADS anxiety sub-score.

    - An improvement in the frequency of symptoms associated with atopic dermatitis in the POEM assessment was seen in both abrocitinib treatment groups compared with the placebo group.

    - An improvement in patient impression of severity of disease within the PtGA was seen in both abrocitinib treatment groups compared with the placebo group.

CLINICAL STUDY REPORT SYNOPSIS

- For EQ-5D-5L index value and EQ VAS, the LSM of change from baseline was improved (higher) for both abrocitinib treatment groups compared with the placebo group beginning at Week 2 that were maintained at all scheduled time points, except for EQ VAS at Week 8 in the abrocitinib 100 mg QD group.

- For FACIT-F, the LSM of change from baseline in adult subjects was improved (higher scores) for both abrocitinib treatment groups compared with the placebo group at Week 12.

- For SF-36v2, the LSM of change from baseline in the physical component summary score was improved (higher) in both abrocitinib treatment groups compared with the placebo group at Week 12. The LSM of change from baseline in the mental component summary score was improved (higher) in the abrocitinib 200 mg QD group only compared with the placebo group at Week 12.

- The LSM of change from baseline in the WPAI:AD (including percent impairment while working, percent overall work impairment, and percent activity impairment) was improved (lower) in both abrocitinib treatment groups compared with the placebo group at Week 12. The LSM of change from baseline in the WPAI:AD (percent work time missed) was not different in both abrocitinib treatment groups compared with the placebo group at Week 12.

- For the CDLQI, EQ-5D-Y, and peds-FACIT-F, because of the low numbers of adolescent subjects, results are difficult to interpret.

Pharmacokinetics:

- Abrocitinib plasma concentrations in moderate-to-severe AD subjects increased in a proportion approximately similar to the 2-fold increase in the dose between the abrocitinib 200 mg QD group and abrocitinib 100 mg QD group.

[redacted]

[redacted]

CLINICAL STUDY REPORT SYNOPSIS

SAFETY

- Abrocitinib was well tolerated. Safety concerns were manageable.
- A higher proportion of subjects with all-causality TEAEs occurred in the abrocitinib groups compared with the placebo group.
    - All-causality TEAEs that appeared more commonly in the abrocitinib groups relative to the placebo group included nausea, headache, nasopharyngitis, upper respiratory tract infection, and acne.
    - Most all-causality TEAEs (>95%) were mild or moderate in severity.
    - The incidence of severe TEAEs was similar between the abrocitinib treatment groups and the placebo group, with most related to atopic dermatitis.
    - TEAEs relating to atopic dermatitis exacerbation were more frequent in the placebo group compared with the abrocitinib treatment groups.
- The incidence of SAEs was low and similar in all treatment groups (1.3%-3.2%).
- The incidence of subjects who permanently discontinued from the study due to TEAEs was lower in the abrocitinib treatment groups compared with the placebo group.
- There was one fatal event of sudden death in the abrocitinib 100 mg QD group, which was assessed as not related to the study drug by the investigator and was confirmed to meet criteria for a cardiovascular event.
- The incidence of subjects with herpes zoster was low with 2/155 in the abrocitinib 200 mg QD group, 0/158 in the abrocitinib 100 mg QD group, and no subject in the placebo group.
- Platelet counts decreased in a dose-dependent manner in both abrocitinib groups with a nadir at Week 4 and returned toward baseline through Week 12 despite continued abrocitinib administration in the majority of subjects. Two (2) subjects in the abrocitinib 200 mg QD group met pre-specified monitoring criteria, but no subject met pre-specified discontinuation criteria. Median lymphocyte counts, neutrophil counts, and hemoglobin values showed no clinically meaningful median changes.
- There were increases in total cholesterol, LDL-cholesterol, and HDL-cholesterol, for which dose-dependent changes were observed for total cholesterol and HDL-cholesterol. There were no changes over time in LDL/HDL ratio and triglycerides.

CLINICAL STUDY REPORT SYNOPSIS

- Dose-dependent increases in creatine kinase were observed in the abrocitinib groups compared with the placebo group but were not clinically meaningful. There were no TEAEs of rhabdomyolysis.

- There were no malignancies or venous thromboembolic events across treatment groups.

- There were no QTcF interval changes or vital signs changes of critical concern.