ORIGINAL ARTICLE

Characterizing the relationships between patient-reported outcomes and clinician assessments of alopecia areata in a phase 2a randomized trial of ritlecitinib and brepocitinib

R. Winnette,[1,*] A. Banerjee,[2] V. Sikirica,[3,†] E. Peeva,[4] K. Wyrwich[1†]

[1]Patient-Centered Outcomes Assessment, Pfizer, New York, NY, USA
[2]Early Clinical Development, Pfizer, Cambridge, MA, USA
[3]Patient & Health Impact, Pfizer, Collegeville, PA, USA
[4]Inflammation & Immunology, Pfizer, Cambridge, MA, USA
*Correspondence: R. Winnette. E-mail: randall.winnette@pfizer.com

Abstract

Background The phase 2a ALLEGRO trial (NCT02974868) investigated the safety and efficacy of ritlecitinib (PF-06651600) and brepocitinib (PF-06700841) in adults with alopecia areata. No randomized controlled trial for alopecia areata has evaluated correlations between clinician-assessed hair loss and patient-reported outcomes.

Objectives Report scores from the Alopecia Areata Symptom Impact Scale (AASIS; a patient-reported outcome tool) and explore the relationships of those scores with clinician-assessed Severity of Alopecia Tool (SALT) scores at baseline and week 24 of the ALLEGRO trial.

Methods Adults with alopecia areata were randomized to ritlecitinib ($n = 48$), brepocitinib ($n = 47$) or placebo ($n = 47$). After 24 weeks, the mixed-effects model with repeated measures was used to calculate the active treatment groups' AASIS score least-squares mean differences. Relationships between AASIS and SALT scores at baseline and week 24 were evaluated by Pearson's correlation coefficients using pooled data.

Results Baseline AASIS and SALT scores were similar among treatment groups. Both active treatment groups reported significant improvements in AASIS scores at week 24 (least-squares mean differences vs. placebo for ritlecitinib, −0.8 to −2.3; brepocitinib, −0.9 to −3.7; $P < 0.05$ for all). At week 24, the mean SALT scores (standard deviation) improved compared with baseline [ritlecitinib, 54.4 (40.3) vs. 89.4 (15.8); brepocitinib, 31.9 (35.7) vs. 86.4 (18.1)]. The correlation coefficients between AASIS global and subscale scores and SALT scores at week 24 ranged from 0.34 to 0.58; $P < 0.05$ for all.

Conclusions Patients randomized to ritlecitinib or brepocitinib reported significantly improved AASIS and SALT scores at week 24 of the ALLEGRO trial compared to placebo. At week 24, medium-to-large correlations can be seen between AASIS global and subscale scores and SALT scores. Our experience with AASIS instrument highlighted several aspects that suggest new patient-reported outcome tools are needed to accurately assess patients' relevant alopecia areata related signs, symptoms and daily functioning.

Received: 27 April 2021; revised: 29 October 2021; Accepted: 3 December 2021

Introduction

Alopecia areata (AA) is an autoimmune disorder that affects 2% of the global population and causes hair loss ranging from small patches to the entire body.[1–3] The condition is also frequently associated with a significant impact on patients' health-related quality of life (HRQoL) and adverse psychological effects such as anxiety disorders and depression.[4,5] Evaluating these aspects of AA via patient-reported outcome (PRO) tools is crucial in order to more completely understand the effectiveness of treatments.

The phase 2a ALLEGRO trial evaluated the safety and efficacy of two oral Janus kinase (JAK) inhibitors in adults with AA and ≥50% loss of scalp hair.[6] The two JAK inhibitors investigated were ritlecitinib (PF-06651600), an inhibitor of JAK3 and the tyrosine kinase expressed in hepatocellular carcinoma (TEC) family, and brepocitinib (PF-06700841), an inhibitor of tyrosine kinase 2 (TYK2) and JAK1. The primary efficacy endpoint of the trial was the change from baseline in clinician-assessed Severity of Alopecia Tool (SALT) score at week 24. The SALT is a validated instrument for measuring scalp hair loss and divides an individual's scalp into four quadrants: the back, top, left and right sides that represent 24%, 40%, 18% and 18% of the total surface area respectively.[7] The sum of the scores from the four quadrants gives an overall SALT score that ranges from 0% (no loss of scalp hair) to 100% (complete loss of scalp hair).

In addition to changes in SALT scores, the phase 2a ALLEGRO trial also evaluated the change from baseline in Alopecia Areata Symptom Impact Scale (AASIS) scores every 4 weeks up to week 24 as an exploratory endpoint. The AASIS is a PRO tool developed by a team with the stated goal 'to develop a measure of quality of life, symptoms and their impact for patients with AA'.[8] To complete the AASIS, individuals with AA rate each of the 13 items using 'the past week' recall period. There are 7 items assessing signs/symptoms that are rated from 0 (Not present) to 10 (As bad as you can imagine), as well as 6 items assessing interference with daily functioning that are rated from 0 (Did not interfere) to 10 (Interfered completely).[8] An 'AASIS global score' can be calculated as the mean of all 13 items, a 'symptoms subscale' score calculated as the mean of the 7 AASIS signs/symptoms item scores, and an 'interference subscale' score calculated as the mean of the 6 daily functioning item scores, with all mean scores ranging from 0 (best health) to 10 (worst health). The 7 signs/symptoms item scores can also be used to generate 2 other subscale scores: 1 based on the mean of the 2 items evaluating hair loss ('hair loss subscale') and another based on the mean of the remaining 5 items that evaluate other symptoms ('other symptoms subscale'; Fig. 1).

The aim of this analysis was to report the results obtained from the use of the AASIS instrument in the phase 2a ALLEGRO trial and explore the relationships between AASIS scores and clinician-assessed SALT scores at baseline and week 24.

Materials and methods

Patients and assessments

The phase 2a ALLEGRO trial (NCT02974868) was a randomized, double-blind, placebo-controlled, multicentre study with two long-term extensions.[6] The study was conducted at sites in the USA, Canada and Australia and enrolled patients aged 18–75 years with a current episode of fixed hair loss due to AA of ≤7 years duration, involving ≥50% loss of scalp hair and with no spontaneous regrowth during the previous 6 months. Using
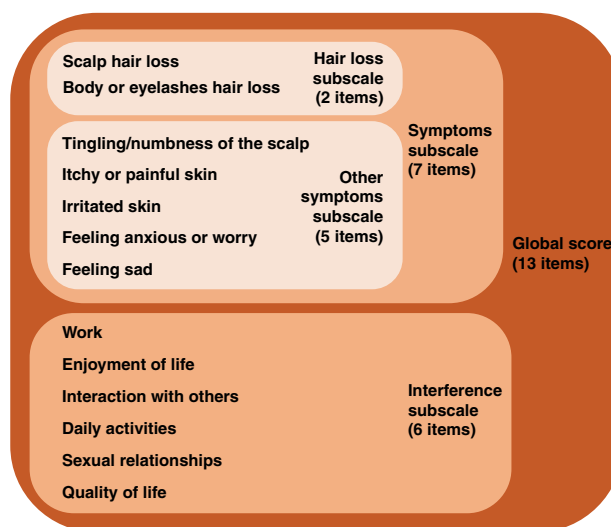

Figure 1 13 items and associated subscales evaluated by the Alopecia Areata Symptom Impact Scale (AASIS) instrument. A patient completing the AASIS instrument rates their experience during 'the past week' on a scale ranging from 0 (Not present) to 10 (As bad as you can imagine) for the 7 signs/symptoms items, and on a scale ranging from 0 (Did not interfere) to 10 (Interfered completely) for the 6 daily functioning items.

interactive response technology, enrolled patients were randomized 2:1:2:1 to receive, respectively, ritlecitinib, ritlecitinib-matching placebo, brepocitinib or brepocitinib-matching placebo. Dosing of ritlecitinib was 200 mg once daily (QD) for 4 weeks followed by 50 mg QD for 20 weeks, and dosing of brepocitinib was 60 mg QD for 4 weeks followed by 30 mg QD for 20 weeks. Data from patients randomized to the two placebo groups were pooled for this analysis. Investigator clinicians evaluated SALT scores at baseline and at weeks 2, 4, 6, 8, 12, 16, 20 and 24. Patients completed the AASIS questionnaire at baseline and at weeks 4, 8, 12, 16, 20 and 24.

Evaluation of SALT scores

Severity of Alopecia Tool scores were previously evaluated as per the objective of the ALLEGRO trial and are described in King *et al.*[6]

Evaluation of AASIS scores and their relationships with SALT scores

Alopecia Areata Symptom Impact Scale scores were evaluated *post hoc* and according to the developer-preferred method, which was published after the statistical analysis plan of the phase 2a ALLEGRO trial was finalized.[8] Briefly, AASIS global and subscale scores were calculated as an arithmetic average of all the items included in each score. Least squares mean differences vs. placebo at week 24 were calculated for the AASIS global score, 4 subscale scores, and 13 individual item scores in the

active treatment groups using a mixed-effects model with repeated measures. Relationships between AASIS and SALT scores at baseline and week 24 were evaluated by Pearson's correlation coefficients (*r* values) using pooled data from all treatment groups. Strength of correlation was categorized according to Cohen's standards: small (*r* = 0.10–0.29), medium (*r* = 0.30–0.49) or large (*r* ≥ 0.50).[9,10] In addition to the correlation coefficient, the respective p-value, and coefficient of determination, ie *r*², are reported. The coefficient of determination (*r*²) details the percentage, expressed as a decimal between 0 and 1, of variation in one item or score that is explained by the other item or score when we examine the correlation. Also, correlation matrix heat maps were generated using R software (version 3.6.1, R Foundation for Statistical Computing, Vienna, Austria) to illustrate the correlation between AASIS global and subscale scores and SALT scores at baseline and at week 24.

Results

Baseline characteristics of the patients

The phase 2a ALLEGRO trial randomized a total of 142 patients with AA whose baseline demographic and disease characteristics have been reported previously.[6] Mean age ranged from 34 to 38 years among the three treatment groups, with the majority gender and race in each group being female (62–77%) and White (77–96%) respectively (Table 1). Mean SALT scores were similar among the treatment groups (86–89% loss of scalp hair), with 42–47% classified as having complete loss of hair on the scalp or complete loss of hair on the scalp and body (alopecia totalis or alopecia universalis respectively; Table 2). In contrast to the mean SALT scores, mean AASIS global scores were low among the treatment groups (2.3–2.8) as were mean scores for the symptoms subscale (2.6–3.4), other symptoms subscale (1.3–2.0) and interference subscale (1.8–2.2). Relative to the mean AASIS global and other subscale scores, the mean hair loss

subscale scores among the three treatment groups were higher (5.9–6.9). The mean scores for the 13 individual AASIS items were similar among the treatment groups.

Effect of JAK inhibitor treatment on AASIS and SALT scores

The mean AASIS global score and SALT score showed a downward trend (improvement) in both active treatment groups but not in the placebo group (Table 2). The mean SALT scores (standard deviation) decreased at week 24 compared with baseline [ritlecitinib, 54.4 (40.3) vs. 89.4 (15.8); brepocitinib, 31.9 (35.7) vs. 86.4 (18.1)]. Both active treatment groups reported statistically significant mean improvements in AASIS global and subscale scores compared with placebo at week 24 (Table 3). In the ritlecitinib group, the AASIS scores for the 2 items in the hair loss subscale, the items 'Itchy or painful skin' and 'Irritated skin' in the other symptoms subscale, and the item 'Sexual relationships' in the interference subscale showed statistically significant mean improvements compared with placebo at week 24. In the brepocitinib group, the AASIS scores for the 2 items in the hair loss subscale, the items 'Feeling anxious or worry' and 'Feeling sad' in the other symptoms subscale, and all 6 items in the interference subscale (Work, Enjoyment of life, Interaction with others, Daily activities, Sexual relationships and Quality of life) showed statistically significant mean improvements compared with placebo at week 24. In both active treatment groups, the items that showed the greatest mean improvements compared with placebo at week 24 were the 2 items in the hair loss subscale [−2.3 for both items in the ritlecitinib group (*P* < 0.01 for both) and −3.7 and −3.5 in the brepocitinib group (*P* < 0.0001 for both)]. Excluding the 2 items in the hair loss subscale, the item that showed the greatest mean improvement compared with placebo at week 24 was 'Itchy or painful skin' in the ritlecitinib group (−1.1; *P* < 0.01) and 'Interaction with others' and 'Sexual relationships' in the brepocitinib group (both −1.6; *P* < 0.001 and *P* < 0.01 respectively).

Correlation between SALT scores and AASIS scores

At baseline, there were small correlations between the SALT score and AASIS Global score (*r* = 0.18, *P* < 0.05) and AASIS symptoms subscale score (*r* = 0.24, *P* < 0.05; Table 4). At week 24, there were large correlations between the SALT scores and the AASIS global score, symptoms subscale score or hair loss subscale score (*r* = 0.51, 0.52 and 0.58; *P* < 0.0001 for all). At week 24, there were also large correlations between the SALT scores and the 2 items in the hair loss subscale (*r* = 0.58 and 0.56; *P* < 0.0001 for both), while the other individual AASIS items had small-to-medium correlations with SALT scores (range of *r* values: 0.23–0.44; *P* < 0.05 for all). Correlation matrix heat map showed positive correlation between AASIS global and subscale scores and SALT scores at baseline (Fig. 2A). However, a stronger positive correlation was observed at week 24 (Fig. 2B).

Table 1 Baseline demographic and disease characteristics

Characteristic†	Ritlecitinib (*N* = 48)	Brepocitinib (*N* = 47)	Placebo (*N* = 47)
Age, years	37 (13)	34 (11)	38 (14)
Female, *n* (%)	37 (77)	32 (68)	29 (62)
White race, *n* (%)	38 (79)	36 (77)	45 (96)
Years since onset of disease, median (range)	6.7 (0.6–52.3)	8.4 (0.3–48.5)	4.8 (0.2–53.4)
Years of current disease episode, median (range)†	2.6 (0.3–7.5)	1.9 (0.2–7.0)	2.4 (0.2–29.5)

Data for age, gender, race, time since onset, duration of current episode are reprinted from King *et al.*[6]
†Data are presented as mean (standard deviation) unless otherwise noted.
‡There were four protocol deviations: three patients in the placebo group and one patient in the ritlecitinib group had a current episode of fixed hair loss >7 years. Three patients (one in each group) had a current episode of fixed hair loss <6 months.

Table 2 SALT and AASIS scores before and after treatment

Characteristic†	Baseline			Week 24		
	Ritlecitinib (N = 48)	Brepocitinib (N = 47)	Placebo (N = 47)	Ritlecitinib (N = 48)	Brepocitinib (N = 47)	Placebo (N = 47)
SALT score†	89.4 (15.8)	86.4 (18.1)	88.4 (18.1)	54.4 (40.3)	31.9 (35.7)	87.4 (20.1)
AASIS global score†	2.8 (1.9)	2.5 (1.9)	2.3 (1.7)	1.8 (2.3)	1.2 (1.8)	2.8 (2.3)
Symptoms subscale score (7 items)	3.4 (1.9)	3.3 (2.0)	2.6 (1.8)	1.9 (2.2)	1.4 (1.8)	2.7 (2.1)
Hair loss subscale score (2 items)	6.9 (3.7)	6.8 (3.8)	5.9 (4.1)	3.7 (4.1)	2.1 (3.2)	5.5 (4.2)
Scalp hair loss	7.4 (3.6)	7.3 (3.8)	6.5 (4.2)	3.9 (4.2)	2.3 (3.4)	5.8 (4.5)
Body or eyelashes hair loss	6.3 (4.3)	6.2 (4.2)	5.4 (4.4)	3.5 (4.1)	2.0 (3.4)	5.2 (4.2)
Other symptoms subscale score (5 items)	2.0 (1.8)	1.9 (1.9)	1.3 (1.6)	1.2 (1.8)	1.1 (1.7)	1.6 (1.7)
Tingling/numbness of the scalp	0.9 (1.9)	0.8 (1.8)	0.9 (1.9)	0.6 (1.8)	0.7 (1.6)	0.9 (1.7)
Itchy or painful skin	2.6 (2.9)	2.4 (2.4)	1.0 (2.0)	0.9 (2.0)	1.4 (2.1)	1.3 (2.2)
Irritated skin	2.0 (2.8)	1.7 (2.6)	1.0 (1.8)	0.7 (1.8)	0.9 (1.6)	1.2 (2.2)
Feeling anxious or worry	2.3 (2.6)	2.7 (2.9)	2.1 (2.8)	2.1 (3.0)	1.3 (2.5)	2.7 (3.0)
Feeling sad	1.9 (2.3)	2.1 (2.7)	1.4 (2.4)	1.7 (2.7)	1.2 (2.6)	2.2 (3.0)
Interference subscale score (6 items)	2.2 (2.5)	1.8 (2.1)	2.0 (2.1)	1.8 (2.6)	1.0 (2.0)	2.9 (3.1)
Work	1.6 (2.7)	1.3 (2.4)	1.6 (3.0)	1.4 (2.6)	0.8 (1.9)	2.5 (3.2)
Enjoyment of life	2.6 (3.1)	2.2 (2.6)	2.6 (2.6)	2.0 (2.9)	1.2 (2.2)	3.1 (3.4)
Interaction with others	2.6 (2.9)	1.9 (2.4)	2.0 (2.5)	2.0 (2.9)	0.9 (2.1)	2.9 (3.3)
Daily activities	2.2 (2.8)	1.6 (2.6)	1.6 (2.2)	1.7 (2.7)	1.0 (2.3)	2.5 (3.2)
Sexual relationships	1.6 (2.7)	1.4 (2.5)	1.6 (2.7)	1.6 (2.6)	0.9 (2.1)	3.0 (3.9)
Quality of life	2.4 (2.9)	2.2 (2.5)	2.5 (2.8)	1.9 (2.8)	1.3 (2.3)	3.2 (3.4)

AASIS, Alopecia Areata Symptom Impact Scale; SALT, Severity of Alopecia Tool.
Data for SALT score are reprinted from King *et al.*[6]
†Possible range: 0–100% (higher scores indicate more severe clinical presentation).
‡Possible range for all AASIS scores: 0–10 (higher scores indicate more severe patent-reported impact).

Table 3 Change in AASIS scores at week 24

AASIS score or item	Ritlecitinib		Brepocitinib	
	LS mean difference vs. placebo (90% CI)	P value	LS mean difference vs. placebo (90% CI)	P value
Global score	−1.1 (−1.6, −0.5)	0.0008	−1.5 (−2.1, −1.0)	<0.0001
Symptoms subscale score	−1.3 (−1.9, −0.7)	0.0002	−1.7 (−2.3, −1.1)	<0.0001
Hair loss subscale score	−2.3 (−3.5, −1.1)	0.0008	−3.7 (−4.9, −2.5)	<0.0001
Scalp hair loss	−2.3 (−3.7, −1.0)	0.0022	−3.7 (−5.0, −2.4)	<0.0001
Body or eyelashes hair loss	−2.3 (−3.5, −1.1)	0.0015	−3.5 (−4.7, −2.3)	<0.0001
Other symptoms subscale score	−0.8 (−1.3, −0.3)	0.0081	−0.9 (−1.4, −0.3)	0.0050
Tingling/numbness of the scalp	−0.3 (−0.9, 0.3)	0.1984	−0.3 (−0.8, 0.3)	0.1855
Itchy or painful skin	−1.1 (−1.8, −0.3)	0.0098	−0.6 (−1.3, 0.1)	0.0732
Irritated skin	−1.0 (−1.7, −0.3)	0.0108	−0.5 (−1.2, 0.1)	0.0890
Feeling anxious or worry	−0.5 (−1.4, 0.3)	0.1546	−1.5 (−2.3, −0.6)	0.0024
Feeling sad	−0.7 (−1.5, 0.2)	0.0906	−1.2 (−2.1, −0.3)	0.0153
Interference subscale score	−0.8 (−1.5, −0.1)	0.0319	−1.4 (−2.1, −0.7)	0.0009
Work	−0.7 (−1.6, 0.3)	0.1214	−1.4 (−2.2, −0.6)	0.0024
Enjoyment of life	−0.7 (−1.5, 0.1)	0.0769	−1.2 (−2.0, −0.5)	0.0042
Interaction with others	−0.8 (−1.7, 0.1)	0.0720	−1.6 (−2.4, −0.8)	0.0007
Daily activities	−0.9 (−1.7, 0.0)	0.0502	−1.3 (−2.1, −0.4)	0.0070
Sexual relationships	−1.0 (−2.0, −0.1)	0.0426	−1.6 (−2.5, −0.7)	0.0025
Quality of life	−0.7 (−1.5, 0.2)	0.1098	−1.3 (−2.0, −0.5)	0.0050

AASIS, Alopecia Areata Symptom Impact Scale; CI, confidence interval; LS, leastsquares.

Table 4 Correlation between SALT scores and AASIS scores

AASIS score or item	Baseline			Week 24		
	r value (95% CI)	P value	r^2 value	r value (95% CI)	P value	r^2 value
Global score	0.18 (0.0119, 0.3325)	0.0359	0.0313	0.51 (0.3602, 0.6327)	<0.0001	0.2592
Symptoms subscale score	0.24 (0.0727, 0.3855)	0.005	0.0553	0.52 (0.3697, 0.6382)	<0.0001	0.2668
Hair loss subscale score	0.31 (0.1562, 0.4552)	0.0002	0.0982	0.58 (0.4503, 0.6922)	<0.0001	0.3412
Scalp hair loss	0.24 (0.0799, 0.3917)	0.0038	0.0586	0.58 (0.4457, 0.6892)	<0.0001	0.3367
Body or eyelashes hair loss	0.34 (0.1904, 0.4817)	<0.0001	0.1186	0.56 (0.4182, 0.6710)	<0.0001	0.3107
Other symptoms subscale score	0.08 (−0.0832, 0.2441)	0.3280	0.0068	0.34 (0.1654, 0.4885)	0.0002	0.1135
Tingling/numbness of the scalp	0.06 (−0.1092, 0.2192)	0.5040	0.0032	0.23 (0.0520, 0.3962)	0.0121	0.0535
Itchy or painful skin	−0.07 (−0.2362, 0.0915)	0.3791	0.0055	0.11 (−0.0720, 0.2867)	0.2335	0.0123
Irritated skin	0.03 (−0.1393, 0.1899)	0.7586	0.0007	0.16 (−0.0183, 0.3354)	0.0776	0.0268
Feeling anxious or worry	0.13 (−0.0326, 0.2912)	0.1150	0.0177	0.38 (0.2182, 0.5291)	<0.0001	0.1478
Feeling sad	0.16 (−0.0006, 0.3202)	0.0509	0.0269	0.34 (0.1681, 0.4906)	0.0002	0.1151
Interference subscale score	0.07 (−0.0966, 0.2314)	0.4127	0.0048	0.44 (0.2769, 0.5737)	<0.0001	0.1911
Work	0.04 (−0.1258, 0.2032)	0.6385	0.0016	0.42 (0.2534, 0.5554)	<0.0001	0.1729
Enjoyment of life	0.01 (−0.1503, 0.1791)	0.8611	0.0002	0.42 (0.2589, 0.5595)	<0.0001	0.1770
Interaction with others	0.11 (−0.0511, 0.2742)	0.1745	0.0131	0.44 (0.2828, 0.5769)	<0.0001	0.1951
Daily activities	0.08 (−0.0826, 0.2446)	0.3246	0.0069	0.38 (0.2085, 0.5217)	<0.0001	0.1412
Sexual relationships	0.01 (−0.1550, 0.1744)	0.9064	0.0001	0.36 (0.1880, 0.5072)	<0.0001	0.1282
Quality of life	0.09 (−0.0789, 0.2481)	0.3035	0.0076	0.44 (0.2821, 0.5764)	<0.0001	0.1945

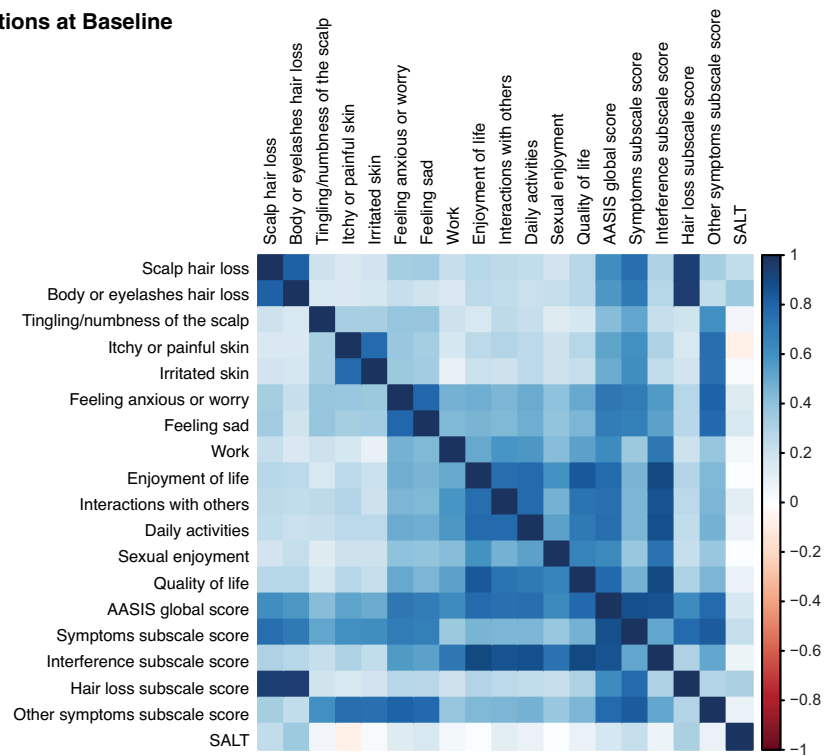AASIS, Alopecia Areata Symptom Impact Scale; SALT, Severity of Alopecia Tool.

Discussion

Our analysis showed that patients with AA who were randomized to the JAK3/TEC inhibitor ritlecitinib or the TYK2/JAK1 inhibitor brepocitinib in the phase 2a ALLEGRO trial reported statistically significant mean improvements in AASIS global and subscale scores and SALT scores after 24 weeks of treatment when compared with patients randomized to placebo. These data reinforce and expand upon previously published results from this trial that showed significant hair regrowth as evaluated by clinician-assessed SALT scores [placebo-adjusted mean decrease in baseline SALT score at week 24 (95% confidence interval): ritlecitinib, 31 (19–44); brepocitinib, 49 (37–62); $P < 0.0001$ for both].[6] To our knowledge, there are only two other interventional trials that have reported AASIS data, but these studies did not report correlations with SALT scores.[11,12]

The mean improvements in AASIS scores observed in the active treatment groups at week 24 ranged from −0.3 to −3.7 on the 0–10 point AASIS scoring system. The AASIS items with the greatest improvement in both active treatment groups were the 2 items in the hair loss subscale ('Scalp hair loss' and 'Body or eyelashes hair loss', mean changes: −2.3 to −3.7), which reflects the significant clinician-assessed hair regrowth previously reported.[6] The remaining 11 items, which are related to symptoms and interference with daily functioning, showed a modest numerical improvement that ranged from −0.3 to −1.6 when compared with placebo at week 24.

While it was hypothesized that the amelioration of hair loss in patients with AA to the extent observed in the ritlecitinib and brepocitinib groups at week 24 of ALLEGRO would be associated with large numerical improvements in all AASIS items, this was not generally observed. This may be due to the low baseline scores for most of the items other than the 2 hair loss items, which provided limited opportunity for numerical improvement. The 11 items relating to symptoms and interference with daily functioning (ie all the items not addressing signs of hair loss) had mean baseline scores that ranged from 0.8 to 2.7 across the 2 active treatment groups despite mean SALT scores in these groups being 86% and 89% at baseline. The observed mean improvement in each of these 11 items at week 24 ranged from −0.3 to −1.6 across the 2 active treatment groups, which represents large proportional decreases relative to baseline but small absolute decreases on the 0–10 scale. This apparent disconnect between high baseline SALT scores and low baseline AASIS scores may reflect issues with the content of the AASIS (eg the items may not address aspects of AA that patients feel significantly impact their lives), the design of the items (eg the wording of the items may be unclear or the response scale inappropriate), patient reluctance to report the extent of their day-to-day challenges due to AA, the characteristics of this specific cohort, or a combination of these factors. The median duration of the current episode of hair loss in the cohort enrolled in the phase 2a ALLEGRO trial ranged from 1.9 to 2.6 years and the median time since disease onset ranged from 4.8 to 8.4 years. Therefore, this cohort was generally comprised of patients who had been living with AA for a considerable length of time, and it is possible that the baseline AASIS scores reflect patient acceptance, coping and/or adaptation to the condition.
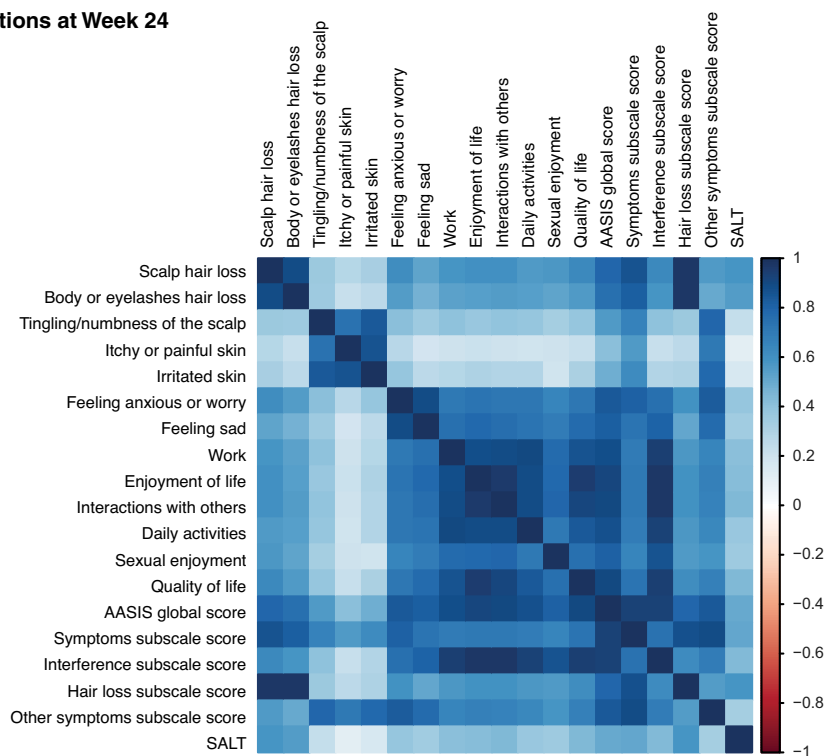
Figure 2 Correlation matrix heat map. Matrices depicting the correlation of Alopecia Areata Symptom Impact Scale global and subscale scores and Severity of Alopecia Tool scores (a) at baseline and (b) at week 24.

Following a retrospective inspection of the AASIS instrument, some limitations of this PRO tool were identified that may have influenced the results of our study. First, some of the items in the symptoms subscale combine two different concepts (a characteristic referred to as being 'double-barrelled'), such as 'Itchy or painful skin', and this may have led to uncertainty regarding response selection if a patient did not experience both concepts. Second, the wording used for the 2 items in the hair loss subscale may be difficult to interpret for patients who have lost all their hair. The AASIS instructions ask the patient to (additional emphasis indicated in bold italics here) 'Please rate how severe the symptoms of your alopecia areata have been *in the past week*. Please select one response from 0 (symptom has not been present) to 10 (the symptom was as bad as you imagine it could be) for each item'.[8] It would be anticipated that patients who have lost all their scalp hair would assign a score of 10 for the 'Scalp hair loss' item, but some may have misinterpreted the intent of this item and assigned a score of 0 because they did not experience any active hair loss/shedding during the previous week. In fact, there were 9 of the 142 patients (6%) in the phase 2a ALLEGRO trial who had a baseline SALT score of 100% (ie complete loss of scalp hair) and assigned a score of 0 for the item assessing scalp hair loss.

The ambiguity in the wordings of some of the AASIS items likely reflects the content-development process for this instrument, which did not include direct, qualitative patient input as recommended by current regulatory guidance on the development of PRO tools for use in clinical trial programmes.[13] The item-generation process used to develop the AASIS involved analyses of patient responses to 125 HRQoL items in the National Alopecia Areata Registry (eg Skindex-16, Dermatology Life Quality Index and Brief Fear of Negative Evaluation Scale)[8,14] and expert clinician review of these items. In the absence of qualitative patient input, the AASIS items derived from these HRQoL items may address concepts that are not salient to the experience of patients with AA. Conversely, these HRQoL items may not address concepts that are a concern to patients with AA in terms of disease-specific physical, emotional and/or social functioning aspects. Indeed, a recent qualitative research study involving direct input from patients with AA demonstrated a lack of relevance for many of the concepts included in the AASIS,[15] and these findings are reinforced by the low baseline scores observed in the cohort that participated in the phase 2a ALLEGRO trial.

Our analysis has some limitations. First, clinical experience with the AASIS instrument and the interpretation of AASIS scores is limited. Meaningful within-patient change thresholds for AASIS scores have not yet been established, and therefore we cannot interpret the clinical meaningfulness of the changes we observed beyond the reported levels of statistical significance.[13] Second, and as discussed above, the low baseline AASIS scores for the 11 items not relating to signs of hair loss (ie those relating to symptoms and interference with daily functioning) in this cohort make it difficult to assess whether these aspects of AA

were meaningfully improved by treatment. Third, while we hypothesized *a priori* that improvement in scalp hair loss, as measured by the SALT score, would yield improvement in the AASIS items by week 24, this assumption deserves closer examination in order to understand the best time frame for measurable patient-reported improvements following regrowth of scalp hair. Finally, this was an analysis of data from a small sample of patients with AA, primarily female and White, who participated in a well-controlled, phase 2a randomized trial, and therefore caution is needed when generalizing our results to broader application or real-world outcomes.

In conclusion, in the phase 2a ALLEGRO trial, patients randomized to the JAK3/TEC inhibitor ritlecitinib or the TYK2/JAK1 inhibitor brepocitinib reported significantly improved AASIS global and subscale scores and SALT scores at week 24 compared with patients randomized to placebo. At week 24, medium-to-large correlations can be seen between the AASIS and SALT scores. Our experience with the AASIS instrument highlighted several aspects that suggest new PRO tools are needed to accurately assess patients' relevant AA-related signs, symptoms and daily functioning.

Data availability statement

Upon request, and subject to certain criteria, conditions and exceptions (see https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information), Pfizer will provide access to individual de-identified participant data from Pfizer-sponsored global interventional clinical studies conducted for medicines, vaccines and medical devices (i) for indications that have been approved in the US and/or EU or (ii) in programmes that have been terminated (ie development for all indications has been discontinued). Pfizer will also consider requests for the protocol, data dictionary and statistical analysis plan. Data may be requested from Pfizer trials 24 months after study completion. The de-identified participant data will be made available to researchers whose proposals meet the research criteria and other conditions, and for which an exception does not apply, via a secure portal. To gain access, data requestors must enter into a data access agreement with Pfizer.

References

1 Lee HH, Gwillim E, Patel KR *et al.* Epidemiology of alopecia areata, ophiasis, totalis, and universalis: a systematic review and meta-analysis. *J Am Acad Dermatol* 2020; **82**: 675–682.

2 Zhou C, Li X, Wang C, Zhang J. Alopecia areata: an update on etiopathogenesis, diagnosis, and management. *Clin Rev Allergy Immunol* 2021; **61**:403–423. https://doi.org/10.1007/s12016-021-08883-0

3 Wambier CG, King BA. Rethinking the classification of alopecia areata. *J Am Acad Dermatol* 2019; **80**: e45.

4 Rencz F, Gulacsi L, Pentek M, Wikonkal N, Baji P, Brodszky V. Alopecia areata and health-related quality of life: a systematic review and meta-analysis. *Br J Dermatol* 2016; **175**: 561–571.

5 Marahatta S, Agrawal S, Adhikari BR. Psychological impact of alopecia areata. *Dermatol Res Pract* 2020; **2020**: 8879343.

6 King B, Guttman-Yassky E, Peeva E *et al*. A phase 2a randomized, placebo-controlled study to evaluate the efficacy and safety of the oral Janus kinase inhibitors ritlecitinib and brepocitinib in alopecia areata: 24-week results. *J Am Acad Dermatol* 2021; **85**: 379–387.

7 Olsen EA, Canfield D. SALT II: a new take on the Severity of Alopecia Tool (SALT) for determining percentage scalp hair loss. *J Am Acad Dermatol* 2016; **75**: 1268–1270.

8 Mendoza TR, Osei J, Duvic M. The utility and validity of the alopecia areata symptom impact scale in measuring disease-related symptoms and their effect on functioning. *J Invest Dermatol Symp Proceed* 2018; **19**: S41–S46.

9 Kiess HO, Green BA. Statistical Concepts for the Behavioral Sciences. Cambridge University Press, Cambridge, 2019.

10 Aberson CL. Applied Power Analysis for the Behavioral Sciences. Routledge, Taylor & Francis Group, New York, 2019.

11 Lai VWY, Chen G, Sinclair R. Impact of cyclosporin treatment on health-related quality of life of patients with alopecia areata. *J Dermatolog Treat* 2021; **32**: 250–257.

12 Fawzy MM, Abdel Hay R, Mohammed FN, Sayed KS, Ghanem MED, Ezzat M. Trichoscopy as an evaluation method for alopecia areata treatment: a comparative study. *J Cosmetic Dermatol* 2021; **20**: 1827–1836.

13 Gnanasakthy A, Mordin M, Evans E, Doward L, DeMuro C. A review of patient-reported outcome labeling in the United States (2011–2015). *Value Health* 2017; **20**: 420–429.

14 Thadanipon K, Suchonwanit P. Measuring patient quality of life following treatment for alopecia. *Patient Prefer Adherence* 2021; **15**: 1601–1610.

15 Winnette R, Martin S, Harris N, Deal LS. Development of the alopecia areata patient priority outcomes instrument: a qualitative study. *Dermatol Ther* 2021; **11**: 599–613.