

Using qualitative methods to establish the clinically meaningful threshold for treatment success for the Severity of Alopecia Tool (SALT) score

Kathleen W. Wyrwich (✉ kathy.wyrwich@gmail.com)

unaffiliated <https://orcid.org/0000-0002-8851-350X>

Helen Kitchen

Clarivate Analytics Inc

Sarah Knight

Clarivate Analytics Inc

Natalie V.J. Aldhouse

Clarivate Analytics Inc

Jake Macey

Clarivate Analytics Inc

Natasha Mesinkovska

University of California Irvine School of Medicine

Justin M. Ko

Stanford University School of Medicine

Brett A. King

Yale University School of Medicine

Research Article

Keywords: alopecia areata, patient-focused drug development, meaningful change, important difference, qualitative research, hair loss

Posted Date: April 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-386715/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Traditionally, appropriate anchors are used to investigate the amount of change on a clinician-reported outcome assessment that is meaningful to individual patients. However, novel qualitative methods can additionally inform the individual improvement threshold for demonstrating the clinical benefit of new treatments. This study aimed to establish a clinically meaningful threshold for treatment success for the clinician-reported Severity of Alopecia Tool (SALT) score for patients with alopecia areata (AA).

Methods: A purposive sample of 10 dermatologists expert in the treatment of AA, and 30 adult and adolescent patients with AA and a history of $\geq 50\%$ scalp hair loss were recruited. Semi-structured interview questions explored thresholds that represented treatment success to clinicians and patients with AA. Findings were analyzed using thematic methods to identify treatment success thresholds.

Results: Expert clinicians considered a static threshold of 80% (n=5) or 75% (n=3) of the scalp hair as a treatment success. Patient responses ranged from 70 - 90% (median 80% of the scalp hair). Subsequently, queried patients confirmed that achieving SALT score ≤ 20 with treatment would be a success. Reflections on the methodology include: clinician perceptions can be informed by current clinical practice and knowledge of existing treatments and research and clinicians easily identified thresholds through qualitative discussions; patients were able to understand and identify thresholds for improvement less than complete absence of disease.

Conclusions: This qualitative investigation of expert clinicians and patients with AA confirmed that achieving an amount of 80% or more scalp hair (SALT score ≤ 20) was an appropriate individual treatment success threshold indicating clinically meaningful improvement for patients with $\geq 50\%$ scalp hair loss. A qualitative investigation of a quantifiable treatment success is possible through well-designed interview questions for the indication and patient population. Both clinician and patient input plays a critical role in understanding the clinical benefit meaningful to patients.

Introduction

Advancing the science of quality of life (QOL) and related patient-centered outcomes - with a focus on improving the QOL for people everywhere by creating a future in which their perspective is integral in health research, care and policy – requires listening closely to the patient voice. This mission and vision of the International Society for Quality of Life Research (ISOQOL) are essential to meeting the needs of persons with AA in the contemporary treatment landscape of new and emerging therapies for treatment of this auto-immune disease, characterized by scalp, facial and/or body hair loss, with a devastating psychological, social and physical toll on those affected and their families [1-6].

Great advancements over the past two decades led to the development and the dissemination of dermatologic training for the Severity of Alopecia Tool (SALT), a systematic method for clinician assessment of scalp hair loss on a 0 (= no missing scalp hair) to 100 (= 100% missing; no scalp hair)

scale [7, 8]. The 2004 introductory publication detailing the SALT scoring process promoted use of the SALT score's percent change metric to understand response to treatment, with 50% improvement from baseline (i.e., SALT₅₀) noted as acceptable endpoint for trials involving extensive alopecia areata ($\geq 50\%$ scalp hair loss at baseline) and systemic agents [7]. Indeed, two landmark proof-of-concept studies investigating the safety and efficacy of oral Janus kinase (JAK) inhibitors in patients with extensive alopecia areata reported the proportion of subjects with 50% or greater hair regrowth from baseline to end of treatment as the primary endpoint [9] or to define the *strong responder* classification [10].

Currently, there are no regulatory-approved treatments for alopecia areata in the US. As noted by the U.S. Food and Drug Administration (FDA), "an important aspect of medical product development is the demonstration of clinical benefit and how that benefit is measured" [11], with similar expectations from other regulatory authorities for registration studies (e.g., European Medicines Agency, Japan's Pharmaceuticals and Medical Devices Agency). The primary, co-primary, or pre-specified secondary endpoints in registration trials used to support medical product approval and labeling claims as well as other communications of clinical benefit are often clinical outcome assessments (COAs), which include: patient-reported outcome (PRO), clinician-reported outcome (ClinRO), observer-reported outcome (ObsRO), performance outcome (PerfO) tools, as well as certain COAs derived from technologies, such as mobile health technologies [11]. The FDA defines *clinical benefit* as "a positive clinically meaningful effect of an intervention on how an individual feels, functions, or survives" and per the FDA, "the process of selecting or developing a COA for use in a medical product development program depends on having adequately characterized the disease or condition, defined the target context of use, and conceptualized a concept of interest that represents clinical benefit." [11]. Moreover, "when a clinical benefit is demonstrated, a description of that benefit can be provided in the regulators' approved labeling or approved communications of the concept or outcome measured (i.e., the aspect of an individual's clinical, biological, physical, functional state, or experience that the assessment is intended to capture)" [11].

With these key directives for registration trial endpoints, an optimal COA for the primary endpoint in future clinical studies with evidence to support a clinically meaningful change (i.e., clinical benefit) is essential to provide a regulatory-approved treatment to patients with AA. While the recommended 50% SALT score improvement from baseline endpoint had been used to define responders in many AA studies, it posed obvious challenges. First, if a patient with no scalp hair (SALT score 100) enrolled in a treatment study of patients with extensive hair loss and achieved this responder status of 50% improvement over time (i.e., achieved SALT score 50), the patient nonetheless continues to have extensive scalp hair loss after treatment. But more importantly, would this patient consider achieving SALT score 50 status to be a *clinical benefit* with "a positive clinically meaningful effect of an intervention on how an individual feels, functions, or survives"? Additionally, SALT₅₀ assumed that scalp hair regrowth was the most important and meaningful treatment outcome for patients. For all stakeholders, it therefore became critical to understand whether scalp hair regrowth was indeed the most important and meaningful treatment outcome for patients with AA versus hair restoration at other locations (e.g., eyebrows, eyelashes).

Furthermore, if this key treatment outcome concept could be soundly established, what is the best estimate for the improvement needed to achieve a clinical benefit at the individual level?

Estimations of meaningful change thresholds on COAs have traditionally been derived through established quantitative methods with anchor-based methods seen as the 'gold standard' and distribution-based methods such as half standard deviation (0.5SD) and standard error of measurement (SEM) seen as supportive [11-15]. In recent years, qualitative methods have emerged as a complementary endeavor [14] to answer this fundamentally patient-centered question, of 'What is a meaningful change for patients?' and there are several examples of interview studies, clinical trial exit interviews, and Delphi studies that have addressed this question [16-20]. Patient perspectives allow further contextualization of the quantitative metric that is derived from anchor-and distribution-based methods; allowing us to understand *why* the score change is important [21].

This emergent qualitative methodology to explore meaningful improvement has mostly focused on within-patient (i.e., individual level) change and not between-group differences. That these qualitative studies have typically focused on exploring within-patient change thresholds has likely been driven in-part by regulatory need in drug development (i.e., the need to define and understand a responder definition) and in-part by practicalities involved in COA development; patients are typically interviewed individually to understand and report on personal experience and not to report on differences between groups of patients or related to treatment. However, there is also a theoretical justification. This focus on individual-level perception of meaningful differences is congruent with the epistemological foundation of qualitative methods based in grounded theory, such as thematic analysis [22], and the phenomenological interpretative approach which seeks to understand the multiple realities of participants rather than one 'true' reality, and focusing on the perceptions, feelings and lived experiences of participants [23].

This study utilized qualitative interview methodology with expert dermatologists and courageous patients with a history of extensive AA to derive the following insights into: a clearer understanding the disease; conceptualization of the most important clinical need; and a categorization of the SALT score into an Investigator Global Assessment (IGA) that could detect clinically meaningful improvement for patients with extensive AA. This paper details the novel qualitative methods used to collaboratively achieve these important goals to establish a clinically meaningful threshold for treatment success in AA.

Methods

Review of Dermatology Endpoints in Recent Product Approvals

Recent (2015-2017) FDA product approvals for dermatologic conditions were reviewed to better understand the primary endpoints previously considered suitable for labelling. The results from this endpoint review informed the Clinician Interview guide.

Clinician Interviews

The Clinician Interviews were conducted to better understand: the clinical diagnosis, management and treatment of patients with AA, and the AA measurement tools that clinicians are most comfortable and supportive in using. In addition, the Clinician Interviews provided detailed insight on clinician perceptions of: the importance of hair loss in specific locations (e.g., scalp, eyebrows, eyelashes, etc.); clinically meaningful change in AA from the clinicians' perspectives; and the clinical relevance and appropriateness of secondary and exploratory clinical study measures.

US dermatologists were identified by Eli Lilly and Company scientists for their contemporary expertise in the diagnosis and treatment of patients with AA, and were recruited through email invitations that outlined the scope of their participation. One-on-one telephone interviews were conducted by an experienced qualitative interviewer trained in COA development techniques (HK or NVJA); interviews lasted 60 minutes and were conducted using a semi-structured interview script, which offered opportunities to systematically explore topics in depth with each clinician while providing consistency across the interviews.

Using the recommendations that emerged from the Clinician Interviews, a Small Panel of two expert clinicians who participated in the interviews was convened to review the quotes and explanations provided in the Clinician Interview data. By incorporating their clinical expertise in AA, the Small Panel's clinicians finalized the draft COA wordings for review by patients.

Patient Interviews

The Patient Interview study protocol was approved by Western Institutional Review Board (ref. #20171820). The interviews were conducted to solicit open-ended patient input to understand the signs and symptoms of AA, the associated impacts and the thresholds for meaningful change that patients considered a treatment success (concept elicitation). The content validity of newly-developed PRO and ClinRO measures were evaluated and documented during the Patient Interviews (cognitive debriefing). The categories of the newly developed IGA were of particular interest to gain patient insights into the categories that would represent a meaningful change. The learnings from the Clinician Interviews informed the semi-structured Patient Interview guide. One-on-one face-to-face interviews lasting 90 minutes - all conducted by the same trained interviewer (NVJA) – offered the opportunity to systematically explore topics in depth while providing consistency across the interviews; however, time constraints did not allow for all patients to debrief all AA measures.

To recruit a patient sample that reflected the range of clinical and demographic characteristics representative of the AA patient population, purposive sampling was used. Minimal sampling target were used to help recruit patients within key demographic and socioeconomic status subgroups. It was important to include patients who had been treated successfully with JAK inhibitors to understand patient perception of the changes in hair growth as well as to enable comparisons of key concepts and assessment of patients' perceptions of clinically meaningful change with JAK inhibitor naïve patients. Additionally, it was important to understand the key concepts and clinically meaningful improvement from the perspectives of patients with eyebrow and/or eyelash involvement in addition to scalp hair

involvement; therefore, this patient group was purposely oversampled. The patients were recruited from two US clinical sites: University of California-Irvine and Yale University in Connecticut. Specific details describing the Patient Interview study inclusion and exclusion criteria have been published [24].

Coding and Analysis

Background clinical experience (clinicians) and demographic and clinical characteristics (patients) were collected during recruitment and summarized. Interviews were audio-recorded with the permission of the interviewee, transcribed, and then examined via thematic analysis assisted by ATLAS.ti Version 7.5 software for the coding and organization of the interview data [25]. During the coding process, all identifying information (e.g., name, specific location, etc.) was removed from the transcripts. Employing a phenomenological interpretative approach, the thematic analysis sought understandings of participants' multiple realities, focused on the individual interviewee's feelings, perceptions, and lived experiences [23].

The following steps for thematic analysis were followed to explore the open-ended concept elicitation interview data:

1. Familiarizations: The lead analysts read the transcripts to identify overarching ideas.
2. Generating codes: Within the transcripts, descriptive codes were generated and then assigned to interviewee quotes.
3. Searching for themes: Using the descriptive codes, potential themes were collated.
4. Reviewing, defining and naming themes: These themes were then compared and contrasted in order to assess any relationships between them, both within and between participants.
5. Reporting: Key concepts and themes were identified within each interview and across the respective samples (clinicians and patients), and supportive quotes extracted and reported.

Data obtained during the cognitive debriefing review of draft items were subject to framework analysis whereby a pre-defined code list was applied to identify the relevance and appropriateness of item wordings, response options and recall periods. As emergent data were expected in the debriefing/item review discussion, iterative codes were also applied. Data from the cognitive debriefing/item review discussions during the Clinician Interviews were ultimately used to amend the draft IGA, and other secondary and exploratory endpoint measures for review and cognitive debriefing during the Patient Interviews.

Results

The FDA-labeled dermatology endpoint review concluded that in general, ordinal, static investigator global assessment (IGA) scales informed the interpretation of clinical trial results in approved dermatologic conditions (2015–2017). Nearly all of the IGAs used in recent dermatology product labeling had 4 or 5 levels; the IGA levels demonstrated a distinct, non-overlapping and clinically relevant gradations of the specific disease/condition, with the highest level (0 or none) representing true clinical

absence of the disease/condition. A static (versus dynamic) IGA assesses the clinician's current impression of disease severity and does not depend on the clinician's recollection of baseline disease severity, which could introduce recall bias [26]. Interestingly, an assessment at one of the top two levels (0 or 1) with at least a 2-level change from baseline was often required to establish individual patient treatment success.

Clinician Interviews

Ten dermatologists, expert in the diagnosis and treatment of AA from across the US of participated in the qualitative telephone interviews in July 2017. On average, these clinicians had been treating or managing patients with AA for 21.2 years (range: 6–38 years).

The clinicians overwhelmingly described scalp hair loss as the most presenting complaint for patients with AA and the primary sign of AA. All interviewed clinicians assessed patients' scalp hair loss in clinical practice, and used either the SALT to evaluate a patient's change over time or made a visual assessment of whether significant improvement had been achieved since a prior visit. Additionally, two of the clinicians made use of photography to monitor the progress of their patients with AA. All clinicians were familiar with the SALT, and nine of the 10 clinicians had used the SALT in AA clinical studies. All clinician agreed that assessments of regrown hair should include only terminal hair (not vellus hair), although the latter may be an early indicator of eventual terminal hair growth.

When queried about their perspectives on AA "treatment success," clinicians noted several factors influencing their judgment of this goal. The *quantity* of scalp hair growth was described as the primary aim for treatment by all clinicians, with other factors were noted by a smaller number of clinicians, including location/pattern of regrowth and hair density.

When asked to describe the amount/percentage of scalp hair they would consider a treatment success, the predominant response was 80% of the scalp hair (n = 5), followed by 75% (n = 3) and 90% (n = 1) (Fig. 1). The remaining clinician did not report a static scalp hair amount, choosing "at least 50% improvement" as the treatment success response.

Figure 1. Clinician Treatment Success Thresholds

Although patient input on the key construct and the appropriate level for that construct to indicate a treatment success/clinical benefit was still needed, clinicians are the primary reporter of scalp hair loss in clinical practice and research. Therefore, clinicians reviewed and iteratively developed the IGA for AA scalp hair loss, with a focus on ensuring distinct and clinically relevant gradations of scalp hair loss. The iterative process used in these IGA discussion is detailed elsewhere [27], and summarized here. Using a top level of 0 (None) representing the absence of scalp hair loss (SALT score 0), the next level (1 = Limited) included SALT score 1–20, with the SALT score 20 upper bound representing the clinician's most commonly-reported treatment success level (Fig. 1). The fourth level (3 = Severe) of the proposed IGA initiated at SALT score 50, and aligned with the lower limit for extensive scalp hair loss [7, 28].

Consequently, the third level (2 = Moderate; SALT score 21–49) was sandwiched between Limited and Severe. Achieving clinician consensus on the draft IGA's fifth level at the highest end of the extensive scalp hair loss spectrum was a challenge; nonetheless, with careful review of all de-identified Clinician Interview responses by the Small Panel members, a relevant description of the 5th level reflecting the clinicians' learnings from patients in this scalp hair loss category (4 = Very Severe; SALT score 95–100) was created to capture patients with nearly complete or complete scalp hair loss. In due course, the draft IGA was reviewed during the Patient Interviews.

The Clinician Interviews also provided insights into other hair loss locations, such as eyebrow and eyelash hair loss. These insights and newly developed sign/symptom COAs are presented elsewhere [24].

Patient interviews

Thirty patients with a history of $\geq 50\%$ scalp hair loss were interviewed in October 2017. The patients' demographic and clinical characteristics at the time of the interviews are detailed elsewhere [29], and synopsized here. Five of the patient interviewees were adolescents (ages 15–17 years old; 3 females/2 males) and 25 of the patient interviewees were adults (ages 18–72 years old; 14 females/11 males). Nine patients were non-Caucasian (Asian, Black, Other). Sixty percent of the adolescents and 84% of the adult patient interviewees had experienced some eyebrow and/or eyelash hair loss, meeting the recruitment goal (80% overall) to oversample these patients with AA. On average, these patients had been diagnosed with AA for 11.4 years (range: 1–46 years). The most recent clinician-assess SALT scores for these patients ranged of 0-100 (mean = 57.9), reflecting the inclusion of patients who had experienced improvements with treatment (60% were currently or previously treated with JAK inhibitors), and the opportunity to obtain recently-informed understandings hair growth changes related to treatment to further understand clinically meaningful change/clinical benefit.

Concept elicitation

The Patient Interviews commenced with discussions of the signs and symptoms experiences with AA, previous treatments and the impacts that AA had on each patient's everyday life and well-being. These discussions were powerful and informative, and the results directly informed a new conceptual model for AA detailing the sign and symptoms AA, and the physical, emotional, and functional impacts of AA, including stigmatization, relationship and social impacts [6].

Ranking exercise

During concept elicitation discussions, the interviewer noted the signs and symptoms mentioned by the patients, and saturation of physical signs and symptoms was achieved [29]. All 30 patients named scalp hair loss as a key sign/symptom. After elicitation of the signs and symptoms experienced, each patient was asked to rank (first/most, second, third) their most bothersome signs and symptoms of AA. Scalp hair loss was named as the most bothersome sign/symptom by 77% of the sample (100% of adolescents/72% of adults). Four adults (16%) named eyebrow hair loss as the most bothersome

sign/symptom; eyelash, nose and body hair loss each received the most bothersome ranking from one adult patient [29]. The results from this patient ranking exercise confirmed scalp hair loss as the key concept, despite oversampling patients with eyebrow/eyelash hair loss.

Meaningful treatment success

All 30 patients were asked to discuss their ideal treatment experience, including both the amount, quality and the time to achieve the hair growth that they would deem clinically meaningful. When patients were asked to propose the percentage (amount) of scalp hair coverage – short of 100% – that they would need for a treatment to be considered successful, 4 patients were initially unable to answer this question, as they experienced some difficulty in discussing scalp hair coverage in terms of percentages. Of the 26 patients (4 adolescents and 22 adults) who were comfortable answering the question, the majority (n = 20) provided answers within the range of 70–90% scalp hair (median = 80% of scalp hair) (Fig. 2), which was generally similar to the Clinician Interviews results (Fig. 1). Moreover, these results were similar for patients with and without JAK inhibitor treatment experience (median = 75% and 85%, respectively). To understand the ‘why’ behind the treatment success metric, patients were asked how they perceived achievement of their desired threshold could impact them. Patients explained how achieving the reported amount of scalp hair would help to improve their emotional/psychological wellbeing such as increasing their confidence levels, reducing stress, and feeling more comfortable around other people. Some improvements to daily life were also predicted as a result of feeling more comfortable around others, such as being able to work more sociable hours and live a more active lifestyle by attending the gym/swimming pool (Table 1).

Table 1

Example Patient Interview Quotes Describing the Meaningful Impact of Achieving the Desired Scalp Hair Amount

<i>I would say 80–90%, because then you could still wear a hat and go swimming and do things like that and cover it up a little bit. (27-F-A-100-N)</i>
<i>It would probably be more of a confidence booster. (28-M-A-100-N)</i>
<i>Well, it's kind of hard to say. I mean, for the most part, I guess I would change my hours of operation. Instead of working at nighttime, I'd work in the daytime. That's what would change my life at least....being around people, yeah. (01-M-A-100-N)</i>
<i>I think just like not being stressed about it because like right now I feel like I can cover it. I don't have to wear a wig or anything like that but and it still is stressful to like worry about like oh it's like exposed or something like that. [...] I'm guessing if it was 80 percent I would be able to do like different hairstyles [...] Also between the two that would be nice and then just generally like not worrying about it as much. I feel like 80 percent I wouldn't like worry about oh exposed or like my hair what is that going to be like? So I think just in terms of like stress and like variety that would be nice. (20-F-P-60-JAK)</i>
<i>Because it's difficult [currently] to, like I said, go swimming, and you can't—the hardest thing for me is like I mean I used to live such an active lifestyle. So I don't like to go to the gym in a scarf, because it's just so hot and sweaty. And I don't like to go to the pool with a bald head, because everyone like stares at you and then the sun. So I would say that however long it took I would be happy with that to get my regular life back. (27-F-A-100-N)</i>
<i>Patient IDs: Order of interview – Sex – Adult/Pediatric – SALT score – Previous JAKi treatment. For example, patient 28-M-A-100-N was interviewed 28th, is a male adult with 100% hair loss and not previously treated with JAKi.</i>

Figure 2. Patient Treatment Success Thresholds

Patients noted that a treatment would be successful even if the scalp hair grown was not the exact same color, quality or thickness as their hair before AA. In fact, most patients expected their hair may grow back differently.

Cognitive Debriefing

During the cognitive debriefing, input from patients was solicited on the relevance, appropriateness and importance of the draft IGA developed during the Clinician Interviews. All 30 patients confirmed agreement with the proposed IGA measure, and none of the patients suggested any further changes to the IGA wording or response levels. Nine patients were asked about their perception of meaningful change as measured by the draft IGA. All nine respondents noted that achieving the Limited (SALT score 1–20) level after nine months would indicate the treatment was successful with affirming quotes including: “*That would be great. That would be fantastic.*” and “*I think that would be a win if you got to Limited.*”

These results confirmed the content validity of the AA-IGA™ as a ClinRO conceptualizing of the most important clinical need that can reflect and detect clinically meaningful improvement for patients with extensive AA. The final AA-IGA™ is published elsewhere [27]. As noted for the Clinician Interviews, patients

also provided reviews and insights on the clinical relevance and appropriateness of other newly developed AA COAs [24, 30].

Discussion

This systematic qualitative investigation of expert clinicians and patients with AA confirmed that achieving an amount of 80% or more scalp hair (SALT score ≤ 20) was an appropriate treatment success threshold indicating clinically meaningful individual improvement for patients with $\geq 50\%$ scalp hair loss. This estimate (80% of scalp hair) was a consistent summary outcome dermatologists and patients. This qualitative investigation of a quantifiable treatment success threshold was possible through the input provided by clinicians and patients who generously and courageously sharing their perspectives in response to well-designed Clinician and Patient Interview questions to aid the one-on-one discussions with each informant group.

Our commitment to hearing the patient voice required us to first listen closely to expert dermatologists to ensure clinically meaningful measurement and categorization of COA scores. This resulted in a categorization of SALT scores into the AA-IGA™ which represented distinct gradations of AA scalp hair loss severity.

The Patient Interviews confirmed the key concept (scalp hair loss) and provided unique data on the amount of scalp hair that patients with a history of $\geq 50\%$ scalp hair loss could consider a treatment success (median = 80%). As a result, the AA-IGA™ and the SALT score ≤ 20 threshold serves as a patient-informed ClinRO reflecting a clinical benefit for patients with $\geq 50\%$ scalp hair loss at the individual patient level [27].

An obvious limitation of this study was that patient perceptions of successful hair growth may be influenced by cultural and societal factors, and the results of this study based on US informants could not be assumed applicable in other countries/cultures. To begin addressing this concern, similar qualitative interviews were conducted in Japan in 2018-19 with expert dermatologists (n = 7) and patients (n = 15) [31]. Both Japanese informant groups confirmed that scalp hair loss was the most important sign/symptom of AA and the greatest treatment priority, and that achieving $\leq 20\%$ scalp hair loss (AA-IGA™ categories 0 or 1) indicated treatment success for patients with $\geq 50\%$ scalp hair loss. These results increase confidence in this COA and the threshold for achieving a clinically meaningful difference/clinical benefit for patients with AA.

Currently in 2021, SALT score ≤ 20 is the primary endpoint responder definition in several on-going Phase 2 or Phase 3 clinical trial programs for the treatment of AA (e.g., NCT03570749, NCT03732807, NCT04518995). As this threshold for interpreting within-patient meaningful change was derived qualitatively, it is important that the validity of this clinically meaningful threshold also be empirically investigated in the emerging study data.

As experts in what it is like to live with their condition, patients are uniquely positioned to inform the understanding of the therapeutic context for drug development and evaluation [32], and as shown here, informed a ClinRO measure to collect meaningful patient experience data and identify responders according to patients' needs and expectations [33]. Indeed, with this qualitative exploration of a quantitative responder threshold, we can now understand the within-patient change in SALT score change that is empirically meaningful and *why* it is meaningful.

Declarations

Acknowledgments

The authors appreciate the wisdom and courage of each interviewee; the careful recruitment and patient access by clinic site staff and coordinators at the Department of Dermatology, University of California, Irvine, (UCI) and Yale University School of Medicine; and the editorial assistance and review by Emily Heredia and Diane Stothard at Eli Lilly and Company in the preparation of this manuscript.

Study Funding Source: Eli Lilly and Company

Competing interests: Kathleen W. Wyrwich is a stockholder at Eli Lilly and Company and Pfizer Inc. Helen Kitchen, Sarah Knight, Natalie V. J. Aldhouse and Jake Macey are employees and stockholders of Clarivate, a health economic and outcomes research consultancy that consults with various pharmaceutical companies. Natasha Mesinkovska has received honoraria/fees for advisory work with Eli Lilly and Company, Concert Pharmaceuticals Inc, Arena Pharmaceuticals, Nutrafol. Justin M. Ko has served on advisory boards and is a consultant and clinical investigator for Eli Lilly and Company. He has served as a clinical investigator and/or consultant for AbbVie, Sanofi, Regeneron, Dermira, BMS and Arena Pharmaceuticals. He has received consulting fees from Eli Lilly and Company and Arena Pharmaceuticals. Brett A. King has served on advisory boards and/or is a consultant and/or is a clinical trial investigator for Aclaris Therapeutics Inc, Almirall, Arena Pharmaceuticals, Bristol-Meyers Squibb, Concert Pharmaceuticals Inc, Dermavant Sciences Inc, Eli Lilly and Company, Incyte Corp, Pfizer Inc, TWi Biotechnology Inc, and Viela Bio. He is on speaker bureaus for Eli Lilly, Pfizer Inc, Regeneron and Sanofi Genzyme.

Ethics approval and consent to participate: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (Western Institutional Review Board [ref #20171820]) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. 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 Oct;75(4):806-12.e3.

2. 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 Sep;175(3):561-71.
3. Davis DS, Callender VD. Review of quality of life studies in women with alopecia. *Int J Womens Dermatol*. 2018 Mar;4(1):18-22.
4. Wolf JJ, Hudson Baker P. Alopecia Areata: Factors That Impact Children and Adolescents. *Journal of Adolescent Research*. 2019;34(3):282-301.
5. Davey L, Clarke V, Jenkinson E. Living with alopecia areata: an online qualitative survey study. *British Journal of Dermatology*. 2019;180(6):1377-89.
6. Aldhouse NVJ, Kitchen H, Knight S, Macey J, Nunes FP, Dutronc Y, et al. "You lose your hair, what's the big deal?' I was so embarrassed, I was so self-conscious, I was so depressed:" a qualitative interview study to understand the psychosocial burden of alopecia areata. *J Patient Rep Outcomes*. 2020 Sep 11;4(1):76.
7. Olsen EA, Hordinsky MK, Price VH, Roberts JL, Shapiro J, Canfield D, et al. Alopecia areata investigational assessment guidelines—Part II. National Alopecia Areata Foundation. *J Am Acad Dermatol*. 2004 Sep;51(3):440-7.
8. Wambier CG, King BA. Rule of thumb: A simple tool to estimate 1% scalp surface area. *J Am Acad Dermatol*. 2019 Aug;81(2):630-1.
9. Mackay-Wiggan J, Jabbari A, Nguyen N, Cerise JE, Clark C, Ulerio G, et al. Oral ruxolitinib induces hair regrowth in patients with moderate-to-severe alopecia areata. *JCI Insight*. 2016 Sep 22;1(15):e89790.
10. Kennedy Crispin M, Ko JM, Craiglow BG, Li S, Shankar G, Urban JR, et al. Safety and efficacy of the JAK inhibitor tofacitinib citrate in patients with alopecia areata. *JCI Insight*. 2016 Sep 22;1(15):e89776.
11. FDA. Discussion Document for the Patient-Focused Drug Development Public Workshop on 1198 Guidance 3: Select, Develop or Modify Fit-for-Purpose Clinical Outcome Assessments. 1199 Prepared for a Patient-Focused Drug Development Guidance Public Workshop held on 1200 October 15-16, 2018. 2018.
12. McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011 Apr;11(2):163-9.
13. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S. Methods for interpreting change over time in patient-reported outcome measures. *Qual Life Res*. 2013 Apr;22(3):475-83.
14. Coon CD, Cappelleri JC. Interpreting Change in Scores on Patient-Reported Outcome Instruments. *Ther Innov Regul Sci*. 2016 Jan;50(1):22-9.
15. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006 Oct 11;4:79.
16. Gelhorn HL, Kulke MH, O'Dorisio T, Yang QM, Jackson J, Jackson S, et al. Patient-reported Symptom Experiences in Patients With Carcinoid Syndrome After Participation in a Study of Telotristat Etiprate: A Qualitative Interview Approach. *Clin Ther*. 2016 Apr;38(4):759-68.

17. Peay H, Kennedy A, Fischer R, Bronson A, Furlong P. Promoting meaningful clinical trial outcome measures for Duchenne muscular dystrophy. *Neuromuscular Disorders*. 2016;26:S187.
18. McGraw S, Qian Y, Henne J, Jarecki J, Hobby K, Yeh WS. A qualitative study of perceptions of meaningful change in spinal muscular atrophy. *BMC Neurol*. 2017 Apr 4;17(1):68.
19. Arbuckle R, Staunton H, Sully K, Tomkins S, Khindri S, Svedstater H, et al. Use of Both Qualitative and Quantitative Methods to Estimate Meaningful Change Thresholds for Key Endpoints in Pediatric Asthma Trials. *Value Health*. 2019 Mar;22(3):340-7.
20. Kitchen H, Seitz C, Trigg A, Aldhouse N, Willgoss T, Schmitz H, et al. Patients' and clinicians' perspectives on item importance, scoring, and clinically meaningful differences for the Endometriosis Symptom Diary (ESD) and Endometriosis Impact Scale (EIS). *Health Qual Life Outcomes*. 2021 Jan 6;19(1):7.
21. Staunton H, Willgoss T, Nelsen L, Burbridge C, Sully K, Rofail D, et al. An overview of using qualitative techniques to explore and define estimates of clinically important change on clinical outcome assessments. *J Patient Rep Outcomes*. 2019 Mar 4;3(1):16.
22. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology*. 2006 2006/01/01;3(2):77-101.
23. Guest G, MacQueen KM, Namey EE. *Applied thematic analysis*. Sage; 2011.
24. Wyrwich KW, Kitchen H, Knight S, Aldhouse NVJ, Macey J, Nunes FP, et al. Development of Clinician-Reported Outcome (ClinRO) and Patient-Reported Outcome (PRO) Measures for Eyebrow, Eyelash and Nail Assessment in Alopecia Areata. *Am J Clin Dermatol*. 2020 Oct;21(5):725-32.
25. Friese S. ATLAS.ti 7 User Manual Copyright ©2014 by ATLAS.ti Scientific Software Development GmbH, Berlin. All rights reserved. Manual Version: 190.20140902. Updated for program version: 7.5 2014 Contract No.: Document Number|.
26. Langley RG, Feldman SR, Nyirady J, van de Kerkhof P, Papavassilis C. The 5-point Investigator's Global Assessment (IGA) Scale: A modified tool for evaluating plaque psoriasis severity in clinical trials. *J Dermatolog Treat*. 2015 Feb;26(1):23-31.
27. Wyrwich KW, Kitchen H, Knight S, Aldhouse NVJ, Macey J, Nunes FP, et al. The Alopecia Areata Investigator Global Assessment scale: a measure for evaluating clinically meaningful success in clinical trials. *Br J Dermatol*. 2020 Oct;183(4):702-9.
28. Olsen EA. Investigative guidelines for alopecia areata. *Dermatol Ther*. 2011 May-Jun;24(3):311-9.
29. Wyrwich KW, Kitchen H, Knight S, Aldhouse NVJ, Macey J, Nunes F, et al. The Role of Patients in Alopecia Areata Endpoint Development: Understanding Physical Signs and Symptoms. *J Investig Dermatol Symp Proc*. 2020 Nov;20(1):S71-s7.
30. Wyrwich KW, Kitchen H, Knight S, Aldhouse NVJ, Macey J, Nunes FP, et al. Development of the Scalp Hair Assessment PRO™ measure for alopecia areata. *Br J Dermatol*. 2020 Dec;183(6):1065-72.
31. Macey J, Kitchen H, Aldhouse NVJ, Burge RT, Edson-Heredia E, McCollam JS, et al. Dermatologist and Patient Perceptions of Treatment Success in Alopecia Areata and Evaluation of Clinical Outcome Assessments in Japan. *Dermatol Ther (Heidelb)*. 2021 Jan 10.

32. Chalasani M, Vaidya P, Mullin T. Enhancing the incorporation of the patient's voice in drug development and evaluation. Res Involv Engagem. 2018;4:10.

33. FDA. Patient-Focused Drug Development: Collecting Comprehensive and Representative Input. Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. 2018.

Figures

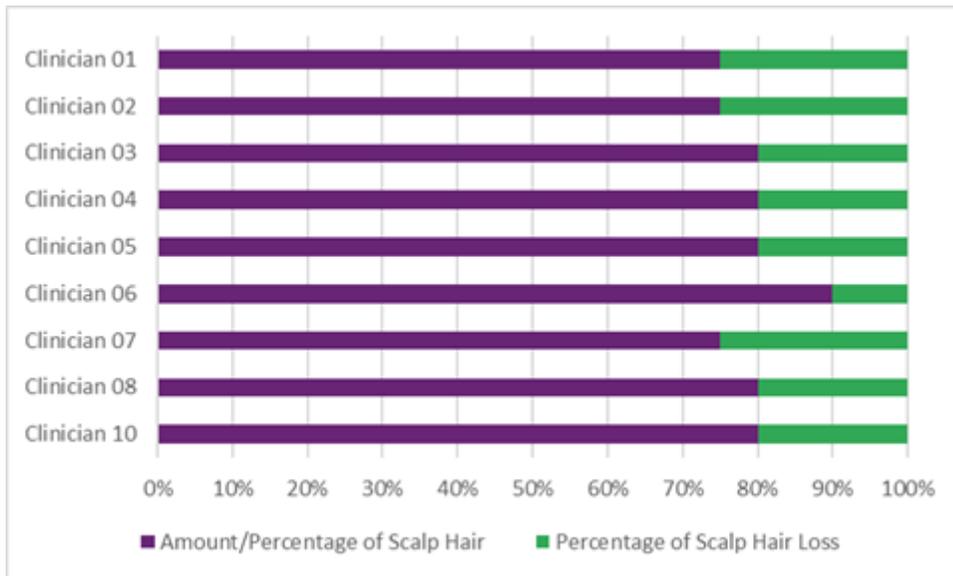


Figure 1

Clinician Treatment Success Thresholds

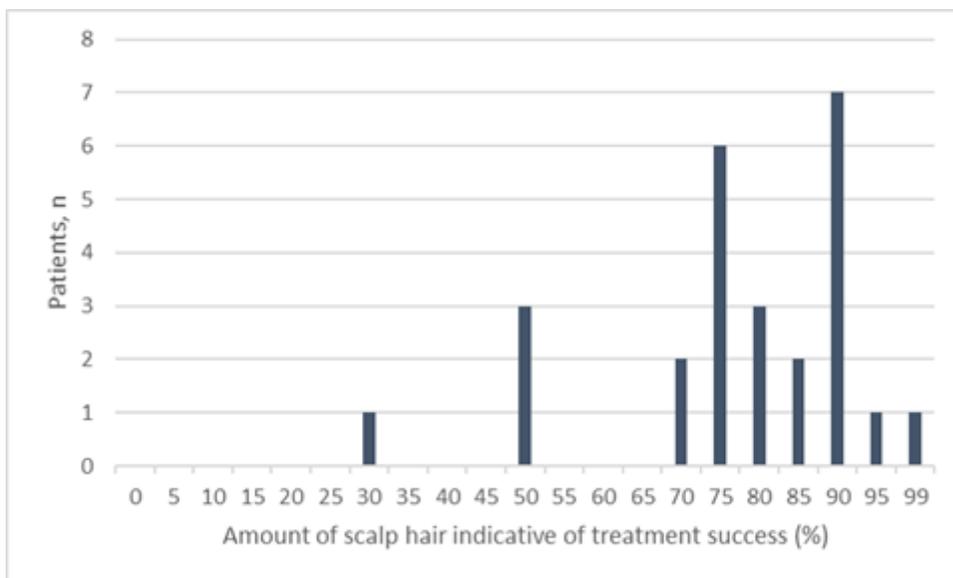


Figure 2

Patient Treatment Success Thresholds