CURRENT OPINION

Expert Perspectives on Key Parameters that Impact Interpretation of Randomized Clinical Trials in Moderate-to-Severe Atopic Dermatitis

Jonathan I. Silverberg[1] · Eric L. Simpson[2] · April W. Armstrong[3] · Marjolein S. de Bruin-Weller[4] · Alan D. Irvine[5] · Kristian Reich[6]

Abstract

The recent advent of numerous clinical trials for the treatment of moderate-to-severe atopic dermatitis has led to new and emerging therapeutic options for this chronic inflammatory skin disease. With this rapid development has come a lack of consistency in study designs, trial conduct, and statistical analyses. Healthcare providers are challenged to interpret how variations in study parameters may influence clinical trial results. Based on literature review and our experience as clinical trialists, we compiled a list of 22 key study parameters of contemporary clinical trials in moderate-to-severe atopic dermatitis and ranked the top study parameters that may have a significant effect on efficacy results. The top parameters included study comparators, rules for rescue treatment, washout periods for topical and systemic treatments, inclusion criteria such as disease severity by Eczema Area and Severity Index and/or Investigator Global Assessment scores, and the duration of the screening period. We describe considerations for these key parameters, with a focus on between-parameter interactions and effect on efficacy results. This may serve to inform the interpretation of atopic dermatitis clinical trials and raise the profile of the need to harmonize the clinical trial design.

Plain Language Summary

Atopic dermatitis (AD) is a skin disease characterized by red, itchy skin that is highly burdensome for patients. Patients with moderate-to-severe disease have large, inflamed skin areas with frequent itching. Recently, the number of clinical trials for drugs that treat moderate-to-severe AD has rapidly increased, with differences in how these trials are designed. There is a need for healthcare providers examining results from different clinical trials to understand how trial design factors might influence study outcomes. In this article, we identify key trial design factors that impact study outcomes, detail how these factors can impact clinical trial results, and explore how these factors interact with one another to affect study outcomes. The five most important design factors, as determined via author surveys, were study comparators (a placebo and/or another drug to which the drug being studied is compared); the rules for the use of rescue treatment (a form of treatment given if an enrolled participant has uncontrolled AD symptoms); washout periods (the time before the trial when previous treatments are stopped to allow them to be cleared from a patient's system); inclusion criteria (that determine which participants are included); and the length of the screening period (the time when patients are assessed to determine if they qualify for participation). By understanding how these key trial design factors impact on study outcomes, healthcare providers may be equipped to better interpret different AD clinical trials. This work also emphasizes the value of harmonizing the AD clinical trial design.

✉ Jonathan I. Silverberg
jonathanisilverberg@gmail.com

1. Department of Dermatology, The George Washington School of Medicine and Health Sciences, 2150 Pennsylvania Ave, NW, Suite 2B-430, Washington, DC 30037, USA

2. Department of Dermatology, Oregon Health and Science University, 3303 S Bond Ave Building 1, 16th Floor, Portland, OR 97239, USA

3. Department of Dermatology, Keck School of Medicine of University of Southern California, 1450 San Pablo St, Los Angeles, CA 90033, USA

4. Department of Dermatology and Allergology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

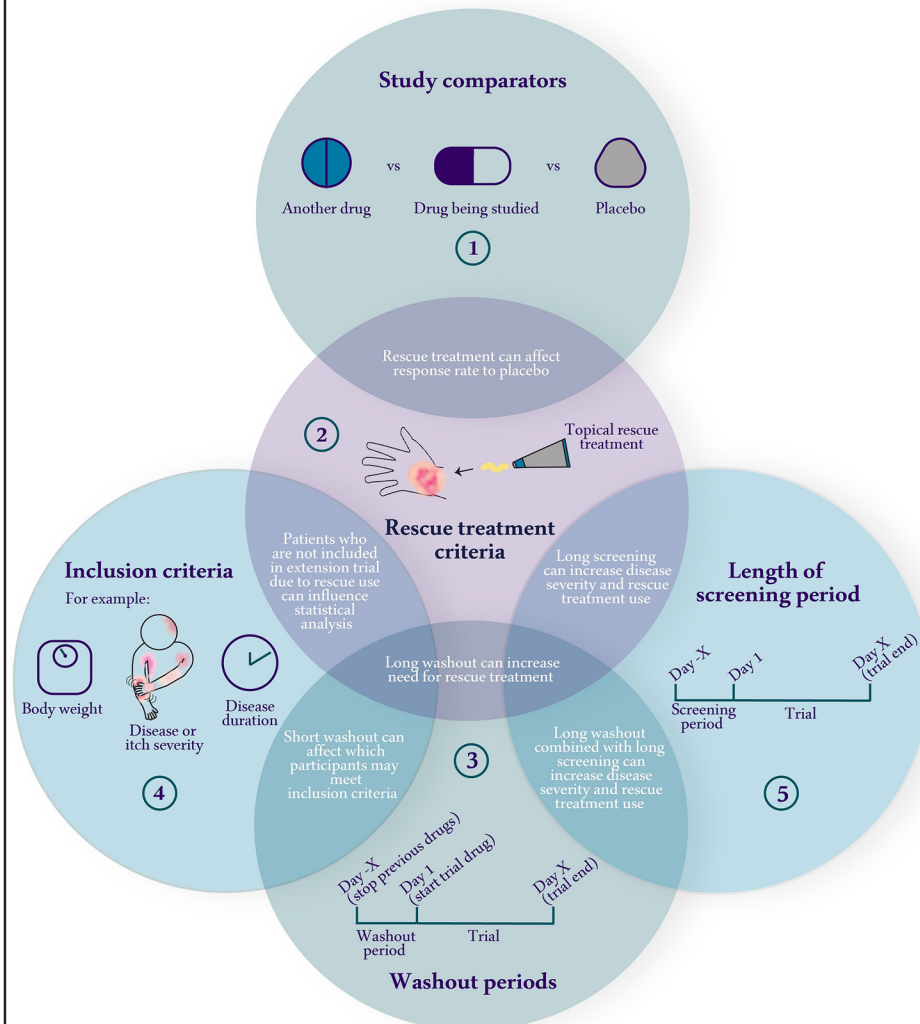5. Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland

6. Translational Research in Inflammatory Skin Diseases, Institute for Health Services Research in Dermatology and Nursing, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

△ Adis

Graphic abstract

Key Points

Contemporary moderate-to-severe atopic dermatitis trials have many parameters that vary and may have significant impact on efficacy results.

Key study design and analysis parameters will need to be considered when interpreting clinical trials results to appropriately implement findings into clinical practice.

1 Introduction

Atopic dermatitis (AD) is a chronic inflammatory skin disease that causes significant morbidity for patients, which when unaddressed may impact psychosocial wellbeing [1–5]. Prior to 2017, treatment options for AD in the US were largely limited to topical corticosteroids (TCS) and off-label use of immunosuppressants [2, 6]. There was recently a marked increase in AD clinical research, particularly for moderate-to-severe disease. The number of clinical trials initiated in participants with moderate-to-severe AD increased from 1 in 2011 to 38 in 2018 (ClinicalTrials.gov, retrieved 23 February 2021). Dupilumab was the first biologic approved for use in patients with moderate-to-severe AD, and another biologic, tralokinumab, has received authorization in the EU and other health authorities [7]. Janus kinase (JAK) inhibitors are also an emerging treatment option [8, 9]. The AD therapeutic pipeline remains robust, with 20 trials initiated in 2020 alone (ClinicalTrials.gov, retrieved 23 February 2021) and 17 trials initiated to date in 2021 (ClinicalTrials.gov, retrieved 24 August 2021). As our understanding of AD pathophysiology evolves, investigations continue to support the development of new treatments through further elucidation of underlying inflammatory pathways and identification of disease biomarkers [6].

As with any rapidly developing area of study, various trial designs have been utilized to assess the effectiveness of investigational treatments for AD [10–21]. Depending on trial timing, guidance from the regulatory bodies may also change, adapting to the evolving treatment landscape. The combination of different trial designs with evolving regulatory guidance poses challenges when interpreting clinical trial results and incorporating evidence into clinical practice.

Clinical dermatologists are left to weigh the importance and impact of trial design factors, which may have a material impact on trial outcomes. While indirect treatment comparisons (e.g., network meta-analyses) are often used to derive some insights from different trials, such efforts remain limited by between-trial differences in key study parameters. There are several ongoing efforts to improve the consistency of clinical research in AD, including the Harmonizing Outcome Measures for Eczema (HOME) initiative, which offers guidance for minimum mandatory outcomes measures [22] and standardizing study design elements in pediatric AD trials [23]. Despite these efforts, important clinical trial design elements remain variable across many phase III studies. As such, those seeking to make evidence-based clinical decisions must evaluate clinical trial results in the context of varying study designs [24], statistical methodology [25], and additional parameters that might impact trial efficacy outcomes.

In this paper, we review which key study parameters may impact efficacy outcomes of clinical trials conducted in participants with moderate-to-severe AD. We also discuss the potential impacts of these parameters on one another. We aim to provide guidance to aid interpretation of randomized controlled trials in AD and raise awareness about the need for harmonization of study design elements in clinical trials conducted in participants with moderate-to-severe AD.

2 Method for Identification of Key Clinical Trial Parameters Impacting Efficacy Outcomes

As a first step in identifying key study parameters in contemporary clinical trials in moderate-to-severe AD, two moderated meetings were convened in December 2020, with participation of all six authors, followed by an initial free-form survey to determine items for consideration. Authors were identified by the sponsor based on their expert knowledge of randomized controlled trials in moderate-to-severe AD; all authors are investigators/authors in recent clinical trials for new systemic and biologic treatments in AD. Survey responses, complemented with findings from a review of the literature and trial protocols for phase III moderate-to-severe AD trials, were used to identify 22 initial trial parameters (Fig. 1). Subsequently, the author group completed a second survey to rank these 22 identified parameters by importance, based on their potential impact on efficacy outcomes. An overall list of ranked parameters was derived from survey responses using the minimum sum of rank scores. Unranked parameters were assigned the average of missing rankings per individual, with the assumption that total rank scores should be the same for all survey respondents (Table 1).
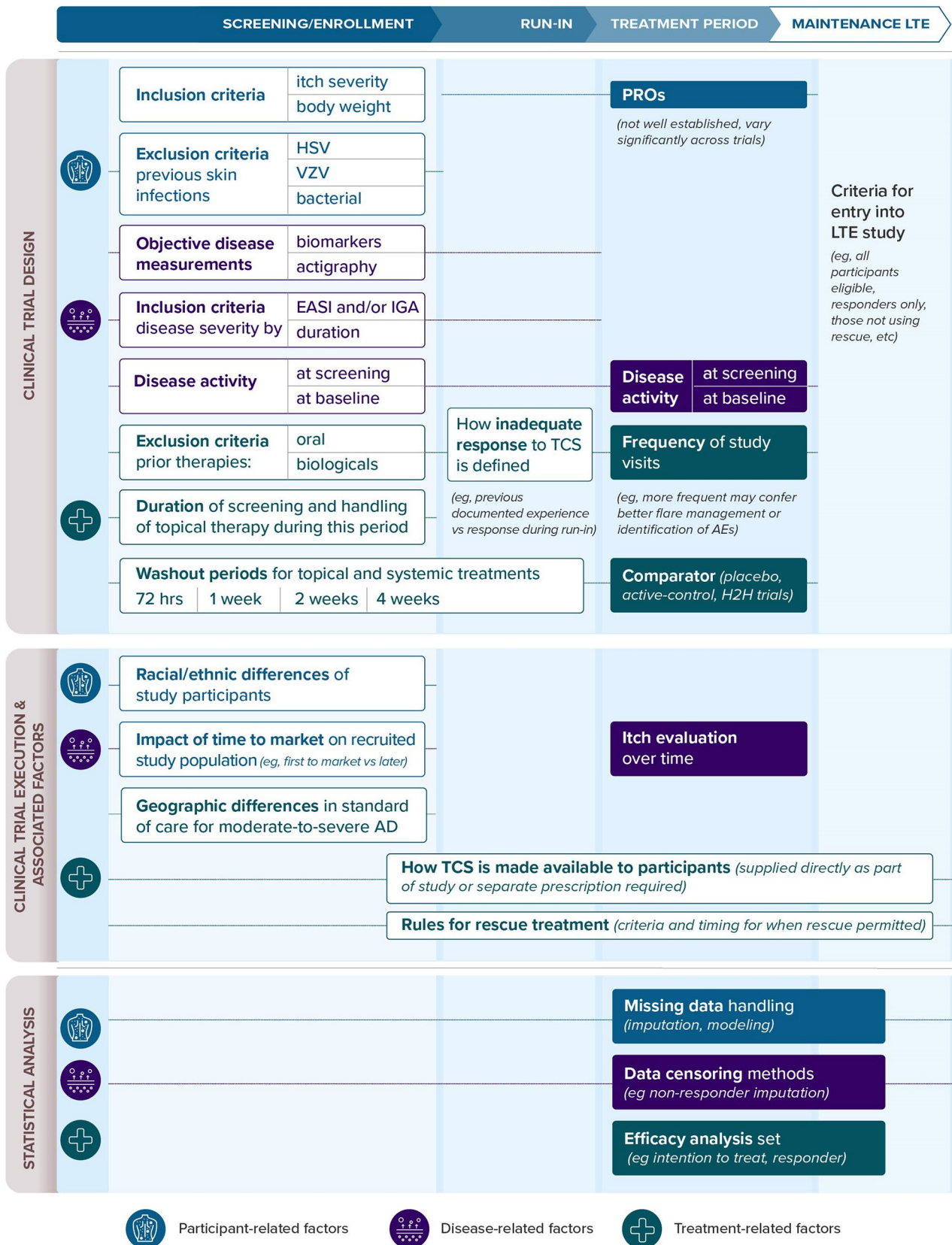
SCREENING/ENROLLMENT ▸ **RUN-IN** ▸ **TREATMENT PERIOD** ▸ **MAINTENANCE LTE**

CLINICAL TRIAL DESIGN

Inclusion criteria
- itch severity
- body weight

Exclusion criteria previous skin infections
- HSV
- VZV
- bacterial

Objective disease measurements
- biomarkers
- actigraphy

Inclusion criteria disease severity by
- EASI and/or IGA
- duration

Disease activity
- at screening
- at baseline

Exclusion criteria prior therapies:
- oral
- biologicals

Duration of screening and handling of topical therapy during this period

Washout periods for topical and systemic treatments
| 72 hrs | 1 week | 2 weeks | 4 weeks |

How inadequate response to TCS is defined
(eg, previous documented experience vs response during run-in)

PROs
(not well established, vary significantly across trials)

| **Disease activity** | at screening |
| | at baseline |

Frequency of study visits
(eg, more frequent may confer better flare management or identification of AEs)

Comparator *(placebo, active-control, H2H trials)*

Criteria for entry into LTE study
(eg, all participants eligible, responders only, those not using rescue, etc)

CLINICAL TRIAL EXECUTION & ASSOCIATED FACTORS

Racial/ethnic differences of study participants

Impact of time to market on recruited study population *(eg, first to market vs later)*

Geographic differences in standard of care for moderate-to-severe AD

Itch evaluation over time

How TCS is made available to participants *(supplied directly as part of study or separate prescription required)*

Rules for rescue treatment *(criteria and timing for when rescue permitted)*

STATISTICAL ANALYSIS

Missing data handling *(imputation, modeling)*

Data censoring methods *(eg non-responder imputation)*

Efficacy analysis set *(eg intention to treat, responder)*

Participant-related factors   Disease-related factors   Treatment-related factors

◄**Fig. 1** Conceptual framework of biologic/systemic therapeutic trials in moderate-to-severe AD. Multiple parameters spanning clinical trial design (top), execution (center), and statistical analysis (bottom) may influence efficacy outcomes over the course of a phase III clinical trial for moderate-to-severe AD. *AD* atopic dermatitis, *AEs* adverse events, *EASI* Eczema Area and Severity Index, *H2H* head-to-head, *HSV* herpes simplex virus, *IGA* Investigator Global Assessment, *LTE* long-term extension, *PROs* patient-reported outcomes, *TCS* topical corticosteroids, *VZV* varicella-zoster virus

3 Key Clinical Trial Parameters that Impact Interpretation of Efficacy Outcomes

3.1 Inclusion/Exclusion Criteria

Inclusion and exclusion criteria are intended to specify an appropriate study population for a given trial. Certain criteria can directly impact observed trial safety/efficacy outcomes and therefore it is important to understand differences in inclusion/exclusion criteria when comparing studies. Often, subtle details may skew the make-up of the study population.

For example, if one clinical trial was to exclude participants with asthma while another included this population, the trial that excludes asthma participants may mitigate any adverse event signal associated with asthma compared with a trial that enabled participants with comorbid asthma to enter. This could similarly be the case for the inclusion/exclusion of participants with prior infections, major adverse cardiovascular events, or venous thrombotic events. Such differences in observed safety outcomes may be more attributable to patient selection than treatment effect.

Efficacy may also be affected if participants in a given trial were not permitted to enter unless they had previously not responded to certain treatments. For example, if the trial only enrolled participants who had previously not responded to a certain standard-of-care treatment, the trial may examine a comparatively more severe and refractory patient population than a study that did not have such inclusion criteria.

3.2 Washout Period Duration

Washout period refers to the time prior to baseline when previous treatments are cleared from a patient's system, such that effects of an investigational drug are not confounded by the previous treatment [26]. Washout period duration varies depending on the medication with generally longer washout periods for systemic therapies relative to topical therapies [10–21]. Optimal washout length also ensures that the patient is at the eligible baseline disease severity.

The length of a washout period may have cascading consequences on various aspects of the subsequent stages of a trial. A longer period (e.g., 4 weeks) reduces the likelihood of a carryover treatment effect, but may result in greater disease severity at baseline, potentially leading to increased rescue treatment use, especially early in a study [27]. Conversely, a shorter washout period may mean the prior treatment still impacts the patient and may reduce the likelihood that participants will flare prior to randomization or may need rescue treatment (typically TCS) to control disease shortly after randomization. Longer washout periods may be a deterrent/disincentive to patient participation, especially among participants with more severe AD due to fear of flaring, raising ethical considerations. Indeed, patient input into study protocol design contributed to the decision to use a short washout period in a recent trial [15]. This relationship between washout length and baseline disease severity warrants further investigation.

3.3 Comparator

Placebo-controlled randomized trials are the gold standard for clinical research and are required for regulatory approval. Considerations around placebo utilization in AD clinical trials and practical suggestions for trial design were recently expertly reviewed [27]. When attempting to compare active agents from independent placebo-controlled trials, attention must be paid to the response rate to placebo in each trial, which may be influenced by background treatments/rescue therapies.

To reflect real-world practice, many clinical trials for moderate-to-severe AD compare placebo and TCS with the experimental drug and TCS, as placebo alone may not meet standard-of-care and may expose participants to a risk of worsening AD and reduced quality of life [27]. However, if background TCS is used on an as-needed basis, participants in the placebo arm may use more TCS than participants in the active treatment arm [10, 13]. This could result in higher levels of AD improvement in the placebo arm and lead to an underestimation of the perceived placebo-adjusted treatment effect of the experimental drug.

Placebo responses may also be influenced by the method by which participants obtain rescue treatment in a trial. Some trials supply TCS directly to participants [14, 17], which minimizes barriers to use. In such trials, participants may use TCS more consistently and in greater amounts than they would in the real-world where a prescription is typically required [28]. Other trials do not supply TCS and require participants to obtain new TCS prescriptions [10, 13]. In some instances, there may be a financial barrier to using TCS if these are not provided. This potentially leads to non-standardized TCS potencies, vehicles, quantities/application frequencies, and likely lower quantities of TCS used overall [29].

Table 1 Ranking of key clinical trial parameters impacting efficacy outcomes

Clinical trial parameters	Rank by minimum sum[a]	No. of top 10 rankings[b]	Mean ± SD	Authors (no. ranked)					
				1 (13)	2 (14)	3 (10)	4 (10)	5 (11)	6 (10)
Comparator (placebo, active-control, head-to-head trials)	**1**	5	4.9 ± 6.4	1	1	5	4	17.5[b]	1
Rules for rescue treatment (criteria and timing for when rescue permitted)	**2**	6	6.2 ± 2.7	8	6	2	5	10	6
Washout periods for topical and systemic treatments (e.g., 72 h, 1, 2, 4 weeks)	**3**	5	6.7 ± 3.1	11	2	6	9	5	7
Inclusion criteria (disease severity by EASI and/or IGA, and duration)	**4**	3	8.3 ± 7.6	2	11	17[b]	1	2	17[b]
Duration of screening and handling of topical therapy during this period	**5**	4	8.5 ± 5.2	12	7	4	17[b]	3	8
Disease activity at screening and at baseline	**6**	4	8.7 ± 7.9	10	19[b]	3	2	1	17[b]
Missing data handling (imputation, linear regression, modeling)	**7**	4	9.0 ± 5.3	6	13	9	17[b]	7	2
Exclusion criteria (prior therapies: oral or biologics)	**8**	4	9.3 ± 7.1	4	19[b]	17[b]	3	4	9
How TCS is made available to participants (supplied directly as part of study or separate prescription required)	**9**	4	9.3 ± 6.7	9	3	1	17[b]	9	17[b]
Efficacy analysis set (e.g. intention to treat, responder)	**10**	4	10.3 ± 5.8	5	8	17[b]	10	17.5[b]	4
Data censoring methods (e.g. non-responder imputation)	**11**	3	11.4 ± 6.2	18.5[b]	14	10	17[b]	6	3
Racial/ethnic differences of trial enrollment	**12**	3	11.9 ± 5.8	7	5	17[b]	8	17.5[b]	17[b]
Geographic differences in standard of care for moderate-to-severe AD	**13**	3	12.7 ± 5.5	18.5[b]	9	7	7	17.5[b]	17[b]
Criteria for entry into long-term extension study (e.g., all participants eligible, responders only, those not using rescue, etc.)	**14**	2	13.1 ± 5.9	18.5[b]	19[b]	8	17[b]	11	5
Inclusion criteria (itch severity, body weight)	15	1	15.1 ± 6.0	3	19[b]	17[b]	17[b]	17.5[b]	17[b]
Exclusion criteria (previous skin infections: HSV, VZV, or bacterial)	16	1	15.2 ± 5.5	18.5[b]	4	17[b]	17[b]	17.5[b]	17[b]
Frequency of study visits (e.g., more frequent may confer better flare management or identification of AEs)	17	1	15.3 ± 3.1	13	10	17[b]	17[b]	17.5[b]	17[b]
Impact of time to market on recruited patient population (e.g., first to market versus later)	18	1	15.8 ± 4.9	18.5[b]	19[b]	17[b]	6	17.5[b]	17[b]
How inadequate response to TCS is defined (e.g., previous documented experience versus response during run-in)	19	1	16.1 ± 4.1	18.5[b]	19[b]	17[b]	17[b]	8	17[b]
Itch evaluation over time	20	0	16.5 ± 3.3	18.5[b]	19[b]	17[b]	17[b]	17.5[b]	10
PROs (not well established, vary significantly across trials)	21	0	16.5 ± 2.3	18.5[b]	12	17[b]	17[b]	17.5[b]	17[b]
Objective disease measurements, such as biomarkers	22	0	17.7 ± 0.9	18.5[b]	19[b]	17[b]	17[b]	17.5[b]	17[b]
Other (please specify)		0	17.7 ± 0.9	18.5[b]	19[b]	17[b]	17[b]	17.5[b]	17[b]

Responses to the second survey varied, as demonstrated by the SDs of rank scores for certain parameters. There was reasonable consensus on the most important parameters, with at least five of the six respondents ranking the following trial design elements within their top 10: rules for rescue treatment, washout periods, and comparator treatment(s). Six parameters focusing on screening duration, disease activity at baseline and screening, treatment history, and statistical analyses were each ranked in the top 10 by four respondents. Fourteen parameters were ranked in the top 10 by at least two respondents, indicated in bold

[a]Scores per comparator were summed and ranked from lowest to highest total value

[b]Unranked parameters were assigned the average of missing rankings per individual. Total rank scores were assumed to be the same for all respondents, regardless of the number of ranked parameters

AD atopic dermatitis, *AEs* adverse events, *EASI* Eczema Area and Severity Index, *HSV* herpes simplex virus, *IGA* Investigator Global Assessment, *PROs* patient-reported outcomes, *SD* standard deviation, *TCS* topical corticosteroids, *VZV* varicella-zoster virus

3.4 Use of Rescue Treatment

As previously mentioned, rescue treatment may be offered to ensure AD is adequately controlled for study participants, with TCS being the most common rescue treatment in AD monotherapy trials. Rescue treatment rules may vary substantially across clinical trials, with some permitting use throughout the active trial periods [10, 13, 17] and others limiting use to a defined period (e.g., only after 2 weeks of treatment) [14] or limiting rescue treatment use entirely [15, 18]. When TCS as needed is used for all study arms, the precise regimen for the rescue treatment may be defined differently across clinical trials. Different trials may also provide varying degrees of patient instruction around rescue treatment use. These variations may drive differences in trial outcomes; the potential efficacy implications of rescue treatment use across the various clinical trial phases are illustrated in Fig. 2.

Given the waxing and waning nature of AD, screening duration and washout specifications during this period can impact disease severity (e.g., Investigator Global Assessment [IGA] or Eczema Area and Severity Index [EASI] scores) at baseline [27]. As previously discussed, participants enrolled in trials prohibiting TCS use during screening or with longer washout periods may experience more disease exacerbation at study initiation than those in trials permitting TCS. Trials involving a large proportion of participants with moderate disease may have comparatively higher efficacy results than those mostly comprising participants with severe disease given the differences in baseline disease severity.

Notably, use of rescue treatment during the initial trial period may impact the patient population that enrolls in the long-term extension phase. Entry into the long-term extension phase of some trials may be restricted to participants who complete the initial treatment period without needing rescue treatment. Conversely, in other trials, participants needing rescue treatment may be eligible for the long-term extension phase.

Rescue treatment use may also influence patient behavior during the initial trial period. If participants are aware that they will not be eligible for the long-term extension after rescue treatment, they may be less likely to use it. In trials where rescue use leads to participants being considered non-responders, this behavior change would translate to impacts on the statistical analysis of the study.

3.5 Nature of the Efficacy Analysis Set

Many aspects of statistical analysis may differ between trials and may influence study outcomes, with a key factor being the efficacy analysis set, defined as the patient population used for the determination of efficacy. During the randomized phase, the efficacy analysis set is often well-defined and approved by regulatory bodies. This may not be the case in long-term extension trials. The estimate of overall treatment response may be conservative if a large proportion of participants were non-compliant or withdrew from the trial. Conversely, a responder-enriched analysis set, comprising participants who showed a strong treatment response, may result in an artificially exaggerated response rate. Including all participants in the long-term extension may yield numerically lower efficacy data than only including those who were responders at the time of the primary endpoint.

3.6 Missing Data Handling and Data Censoring

Approaches for handling missing data have grown more standardized in recent years, with international efforts to arrive at statistical frameworks for use across trials for many therapeutic areas [30]. Until such standards are broadly implemented, it is important to note the statistical methods used when interpreting clinical trials.

Missing data may arise due to study discontinuation or lost to follow-up. The advantages and disadvantages of various common imputation approaches for handling missing data were reviewed previously in the context of psoriasis trials [31]. A key learning for moderate-to-severe AD trials is to consider differences in imputation strategies to understand possible statistical biases and inform more meaningful comparisons [31]. For example, some trials may use the last observation carried forward (LOCF) approach, assigning the value recorded at the patient's last visit to subsequent missed visits; this may overestimate the sustained response or underestimate the improved response the participants may have experienced if they stayed in the trial [32].

Data censoring may also be needed in the context of rescue therapy, if a study specifies that participants needing rescue treatment are to be imputed as non-responders. This conservative approach, known as non-responder imputation (NRI), assigns a value of non-response to the censored data point [31, 32]. When NRI is used, the investigational treatment may be perceived to have lower efficacy, regardless of treatment response up to that point [31, 32]. Varying rules around rescue treatment may impact this aspect of trial design. For example, permitting rescue treatment immediately after randomization may result in more participants being considered non-responders due to rescue use in the early weeks of the trial.

SHORTER
PRIOR TCS WASHOUT
(≤1 WEEK)[15]

May reduce likelihood of rescue
due to short time without treatment

May enroll unstable participants
more likely to have flares and
rescue treatment during early
treatment period phase

LONGER
PRIOR TCS WASHOUT
(>1 WEEK)[10, 12-14, 16, 17, 20, 21]

May result in increased use of TCS
especially in the early weeks of study
period due to participants being
without treatment for longer period

May allow participants longer time to
flare and **meet entry criteria**

SCREENING

Rescue
PROHIBITED[15, 18]

May result in greater rate of **trial
discontinuation**

May result in **fewer participants
being imputed as non-responders**
for studies with NRI analyses

Rescue
PERMITTED[10-14, 16, 17, 19, 21]

May incentivize **greater trial
participation** rates given eventual
access to rescue treatment
(eg, 2 weeks following treatment
initiation), if needed

May result in **more participants
being imputed as non-responders**
for studies with NRI analyses,
lowering the reported response
rates

**TREATMENT
PERIOD**

Rescue-treatment recipients
EXCLUDED

May lower response rate in the
active arm, thereby potentially
leading to lower efficacy if such
participants are considered
non-responders.[15, 18]

Rescue-treatment recipients
INCLUDED

May inflate true response to
treatment in the active arm if
subjects receiving rescue
treatment are not excluded from
the efficacy analyses as
non-responders[10, 13, 14, 17]

**STATISTICAL
ANALYSES**

Rescue-treatment recipients
NON-ELIGIBLE

**May elevate perceived efficacy of
study treatment** if enrollment is
limited to best responders (ie,
those not requiring rescue)

Rescue-treatment recipients
ELIGIBLE

May allow for a broader spectrum
of participants to be enrolled, with
**outcomes more likely to reflect
actual treatment efficacy**

**LONG-TERM
EXTENSION
TRIAL**

Fig. 2 Efficacy implications of variations in rescue treatment. The duration of prior therapy washout during screening may impact the likelihood of flares, which subsequently impacts the need for rescue treatment in the treatment period. Additionally, residual treatment effects due to a short washout period may mask the true severity of the study population, skewing the participants who may meet the inclusion criteria for the treatment period. During the treatment period, whether or not rescue treatment is permitted has implications for trial discontinuation rates. Those who need rescue treatment may be considered non-responders in some trials; how non-responders are statistically accounted for may influence response rates. Long-term extension trials may also set rules around the inclusion/exclusion of participants who required rescue treatment during the treatment period, which may influence the patient profile in this phase. *NRI* non-responder imputation, *TCS* topical corticosteroids

4 Concluding Remarks

4.1 Discussion

There are many ongoing and completed clinical trials for the treatment of moderate-to-severe AD, and their methodological differences make it challenging to comparatively interpret trial results. It is critical that the key study parameters, including comparator, rules for rescue treatment, and washout periods, are recognized and their implications understood. We have provided our perspectives on such parameters, with reasonable consensus regarding those of highest importance. There remains a paucity of evidence in AD to inform clinical trial designs and statistical analysis, and learnings from other therapeutic areas, such as psoriasis, may offer some insights into clinical trial design and interpretation [31].

Increased harmonization of primary outcome measures will also serve to aid the interpretation of AD clinical trials. Outcomes recommended by the HOME initiative could be widely utilized as the minimum set of outcomes, with any additional outcomes dictated by the objectives of a particular study [22]. In addition to type of outcome measurement, the timing of measurements would also be helpful to standardize; an examination of early efficacy in addition to longer-term endpoints at 6 or 12 months to assess sustained AD control would be valuable. However, certain trial design factors, such as study visit schedule or study population, may remain subject to variation due to unique study objectives, differences in drug mechanism of action, or hypotheses being tested.

Another variable aspect of clinical trials that confounds the interpretation of results is participant adherence to the study drug. Reduced adherence to a study drug may lower response rates due to lower effective dose of the drug, which in turn could increase the need for rescue treatment. Such adherence issues are common in the treatment of AD, particularly for topical treatments, with poor adherence in clinical trial settings potentially due to complicated administration instructions, time burden for treatment application, or safety concerns [33, 34]. While drug adherence itself is not a component of trial design, trials may monitor and/or report adherence differently, and low adherence may have cascading impacts on other aspects of a trial in a manner similar to the parameters discussed here. Studies of topical therapies often weigh tubes to measure drug use, and this is commonplace for documenting TCS rescue therapy use in AD trials. However, monitoring adherence in studies of oral or injectable therapies is more challenging; in psoriasis trials, studies using these administration routes often rely on participant reports or prescription refill records, which have variable reliability [35]. The assessment of AD clinical trial results may include considerations of the impact of adherence on results, rescue medication use, and subsequent missing data handling if non-adherent patients are excluded from the results.

In addition to the points raised herein, across-trial interpretation may benefit from increased reporting of the number needed to treat (NNT) and absolute risk reduction (ARR). The NNT represents the number of participants needed to be treated to achieve one additional positive outcome relative to a control group [36, 37]. The ARR expresses the difference in frequency of negative/unfavorable outcomes in the arms of a trial [38]. Since 2001, the CONSORT group (CONsolidated Standards Of Reporting Trials) recommends reporting the NNT in the results of randomized controlled trials [39]. While NNT and AAR also have limitations related to study design, their regular reporting would provide an additional measure when seeking to understand differences between agents studied in independent trials [40].

4.2 Conclusions

Overall, this article highlights the key aspects of clinical trial design that may impact treatment efficacy, and will limit the validity of side-by-side comparisons of clinical trial data. Our review highlights how certain study parameters may impact other aspects of study design. These interactions pose challenges for medical experts seeking to compare the outcomes of two different trials. The differences between trials may impact active treatment and placebo arms differently, which further complicates comparisons. This raises a need for caution towards trial comparisons and substantiates the need to harmonize study procedures in moderate-to-severe AD trials and establish a standardized approach for analyzing trial results. We welcome further research investigating the impact of these differences, as well as efforts to harmonize AD trials.

Declarations

References

1. Barbarot S, et al. Epidemiology of atopic dermatitis in adults: results from an international survey. Allergy. 2018;73(6):1284–93.
2. Langan SM, Irvine AD, Weidinger S. Atopic dermatitis. Lancet. 2020;396(10247):345–60.
3. Silverberg JI, et al. Patient burden and quality of life in atopic dermatitis in US adults: a population-based cross-sectional study. Ann Allergy Asthma Immunol. 2018;121(3):340–7.
4. Simpson EL, et al. Patient burden of moderate to severe atopic dermatitis (AD): insights from a phase 2b clinical trial of dupilumab in adults. J Am Acad Dermatol. 2016;74(3):491–8.
5. Zuberbier T, et al. Patient perspectives on the management of atopic dermatitis. J Allergy Clin Immunol. 2006;118(1):226–32.
6. Paller AS, Kabashima K, Bieber T. Therapeutic pipeline for atopic dermatitis: end of the drought? J Allergy Clin Immunol. 2017;140(3):633–43.
7. European Medicines Agency. Adtralza. 2021 August 19, 2021 [cited 2021 October 21, 2021]; Available from: https://www.ema.europa.eu/en/medicines/human/summaries-opinion/adtralza.
8. Liu A. More relief for Dupixent as FDA pushes back Lilly, Pfizer eczema decisions on safety concerns. 2021 [cited 19 May 2021]. Available at: https://www.fiercepharma.com/marketing/more-relief-for-dupixent-as-fda-pushes-back-lilly-incyte-pfizer-eczema-decisions-amid-jak.
9. European Medicines Agency. Olumiant. 2020 [cited 17 May 2021]. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/olumiant.
10. Simpson EL, et al. Two phase 3 trials of dupilumab versus placebo in atopic dermatitis. N Engl J Med. 2016;375(24):2335–48.
11. Blauvelt A, et al. Long-term management of moderate-to-severe atopic dermatitis with dupilumab and concomitant topical corti-costeroids (LIBERTY AD CHRONOS): a 1-year, randomised, double-blinded, placebo-controlled, phase 3 trial. Lancet. 2017;389(10086):2287–303.
12. Guttman-Yassky E, et al. Efficacy and safety of lebrikizumab, a high-affinity interleukin 13 inhibitor, in adults with moderate to severe atopic dermatitis: a phase 2b randomized clinical trial. JAMA Dermatol. 2020;156(4):411–20.
13. Wollenberg A, et al. Tralokinumab for moderate-to-severe atopic dermatitis: results from two 52-week, randomized, double-blind, multicentre, placebo-controlled phase III trials (ECZTRA 1 and ECZTRA 2). Br J Dermatol. 2021;184(3):437–49.
14. Reich K, et al. Efficacy and safety of baricitinib combined with topical corticosteroids for treatment of moderate to severe atopic dermatitis: a randomized clinical trial. JAMA Dermatol. 2020;156(12):1333–43.
15. Silverberg JI, et al. Efficacy and safety of abrocitinib in patients with moderate-to-severe atopic dermatitis: a randomized clinical trial. JAMA Dermatol. 2020;156(8):863–73.
16. Simpson EL, et al. Baricitinib in patients with moderate-to-severe atopic dermatitis: results from a randomized monotherapy phase 3 trial in the United States and Canada (BREEZE-AD5). J Am Acad Dermatol. 2021;85(1):62–70.
17. Simpson EL, et al. Baricitinib in patients with moderate-to-severe atopic dermatitis and inadequate response to topical corticoster-oids: results from two randomized monotherapy phase III trials. Br J Dermatol. 2020;183(2):242–55.
18. Simpson EL, et al. Efficacy and safety of abrocitinib in adults and adolescents with moderate-to-severe atopic dermatitis (JADE MONO-1): a multicentre, double-blind, randomised, placebo-controlled, phase 3 trial. Lancet. 2020;396(10246):255–66.
19. Silverberg JI, et al. Tralokinumab plus topical corticosteroids for the treatment of moderate-to-severe atopic dermatitis: results from the double-blind, randomized, multicentre, placebo-controlled phase III ECZTRA 3 trial. Br J Dermatol. 2021;184(3):450–63.
20. Guttman-Yassky E, et al. Once-daily upadacitinib versus placebo in adolescents and adults with moderate-to-severe atopic derma-titis (Measure Up 1 and Measure Up 2): results from two repli-cate double-blind, randomised controlled phase 3 trials. Lancet. 2021;397(10290):2151–68.
21. Reich K, et al. Safety and efficacy of upadacitinib in combina-tion with topical corticosteroids in adolescents and adults with moderate-to-severe atopic dermatitis (AD Up): results from a ran-domised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2021;397(10290):2169–81.
22. Chalmers JR, et al. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). Br J Dermatol. 2018;178(5):e332–41.

23. Siegfried EC, et al. Optimizing clinical trials for atopic dermatitis in children. J Invest Dermatol. 2017;137(6):1363–4.

24. Silverberg JI. Study designs in dermatology: a review for the clinical dermatologist. J Am Acad Dermatol. 2015;73(5):721–31 (**quiz 731–2**).

25. Silverberg JI. Study designs in dermatology: practical applications of study designs and their statistics in dermatology. J Am Acad Dermatol. 2015;73(5):733–40 (**quiz 741–2**).

26. Silverberg JI. Trial washout periods are methodologically necessary. Am J Manag Care. 2021. https://www.ajmc.com/view/dr-jonathan-silverberg-trial-washout-periods-are-methodologically-necessary.

27. Leshem YA, et al. Optimization of placebo use in clinical trials with systemic treatments for atopic dermatitis: an International Eczema Council survey-based position statement. J Eur Acad Dermatol Venereol. 2019;33(5):807–15.

28. Tier HL, et al. Tolerability of and adherence to topical treatments in atopic dermatitis: a narrative review. Dermatol Therapy. 2021;11(2):415–31.

29. Storm A, et al. One in 3 prescriptions are never redeemed: primary nonadherence in an outpatient clinic. J Am Acad Dermatol. 2008;59(1):27–33.

30. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2019. https://www.gmp-compliance.org/files/guidemgr/E9-R1_Step4_Guideline_2019_1203.pdf.

31. Langley RG, Reich K. The interpretation of long-term trials of biologic treatments for psoriasis: trial designs and the choices of statistical analyses affect ability to compare outcomes across trials. Br J Dermatol. 2013;169(6):1198–206.

32. Papp KA, et al. Analytical approaches to reporting long-term clinical trial data. Curr Med Res Opin. 2008;24(7):2001–8.

33. Feldman SR, et al. The challenge of managing atopic dermatitis in the United States. Am Health Drug Benefits. 2019;12(2):83–93.

34. Patel N, Feldman SR. Adherence in atopic dermatitis. Adv Exp Med Biol. 2017;1027:139–59.

35. Aleshaki JS, et al. Adherence and resource use among psoriasis patients treated with biologics. Expert Rev Pharmacoecon Outcomes Res. 2018;18(6):609–17.

36. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318(26):1728–33.

37. Citrome L. Show me the evidence: using number needed to treat. South Med J. 2007;100(9):881–4.

38. Irwig L, Irwig J, Trevena L, Sweet M. Smart health choices: making sense of health advice. Chapter 18, relative risk, relative and absolute risk reduction, number needed to treat and confidence intervals. London: Hammersmith Press; 2008.

39. Moher D, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2010;63(8):e1-37.

40. Ludwig L, Darmon P, Guerci B. Computing and interpreting the number needed to treat for cardiovascular outcomes trials: perspective on GLP-1 RA and SGLT-2i therapies. Cardiovasc Diabetol. 2020;19(1):65.