The Best Emollient for Eczema (BEE) trial: a randomised trial comparing the effectiveness of four types of commonly prescribed emollients for children with eczema

Statistical Analysis Plan

Trial registration: ISRCTN84540529
Version 1.2 (16 October 2020)
Based on protocol v7.0 (19 November 2019)

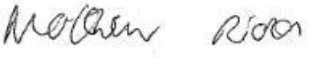
	NAME	TITLE	SIGNATURE	DATE
Author	Stephanie MacNeill	Lecturer in medical statistics		16/10/2020
Study Chief Investigator	Dr Matthew Ridd	Chief investigator		29/10/2020
Statistical reviewer	Dr Ben Carter	Lecturer in medical statistics		16/10/2020

Table of contents

List of abbreviations **4**
1. INTRODUCTION AND PURPOSE **5**
2. SYNOPSIS OF STUDY DESIGN AND PROCEDURES **5**
2.1 Trial objectives and aims **5**
2.1.1 Primary objectives 5
2.1.2 Secondary objectives 5
2.2 Trial design **6**
2.3 Trial centres **6**
2.4 Trial eligibility criteria **6**
2.4.1 Trial inclusion criteria 6
2.4.2 Trial exclusion criteria 6
2.5 Description of interventions **7**
2.6 Randomisation procedures **7**
2.7 Sample size and justification **7**
2.8 Masking **8**
2.9 Trial committees **9**
2.10 Outcome measures **9**
2.10.1 Primary outcome 9
2.10.2 Secondary outcomes 9
2.11 Interim analysis **10**
3. GENERAL ANALYSIS CONSIDERATIONS **10**
3.1 Analysis populations **10**
3.2 Derived variables **10**
3.3 Procedures for missing data **11**
3.3.1 Missing items in questionnaires 11
3.3.2 Missing baseline data 12
3.3.3 Missing outcome data 12
3.4 Study centre effects **12**
3.5 Questionnaire completion **12**
3.6 Timing of final analysis **12**
4. DESCRIPTION OF PARTICIPANT CHARACTERISTICS **13**
4.1 Disposition **13**
4.2 Baseline characteristics **13**
5. ASSESSMENT OF STUDY QUALITY **13**

5.1 Eligibility checks 13
5.2 Data management and data validation 13
5.3 Adherence and contamination 13
5.4 Protocol deviations 14
6. ANALYSIS OF EFFECTIVENESS 14
6.1 Summary of primary and secondary outcomes 14
6.2 Primary analysis 15
6.3 Sensitivity analyses of the primary outcome 15
6.3.1 Missing data 16
6.3.2 Imbalance between treatment groups 16
6.3.3 Self-completed and CSO-collected POEM symptom data at 16-weeks 16
6.3.4 Per protocol analysis 16
6.3.5 Randomisation of ineligible participants 16
6.4 Impact of COVID-19 16
6.5 Sub-group analyses 17
6.6 Secondary outcomes 17
6.6.1 Sensitivity analysis of the EASI score outcome 18
6.7 Characteristics of non-study patients 18
7. ANALYSIS OF SAFETY 18
8. CHANGES MADE TO THE PLANNED STATISTICAL ANALYSES 18
9. BIBLIOGRAPHY 19
APPENDIX A: SKELETON TABLES AND FIGURES 21

List of abbreviations

Acronym	Details
ADQoL	Atopic Dermatitis Quality of Life
BEE	Best Emollient for Eczema
BRTC	Bristol Randomised Trials Collaboration
CCG	Clinical Commissioning Group
CHU-9D	Child Health Utility 9D
CI	Confidence interval
C-I	Chief Investigator
COMET	Choice of Moisturiser for Eczema Treatment
CONSORT	Consolidated Standards of Reporting Trials
CSO	Clinical science officer
DFI	Dermatitis Family Impact
DMC	Data Monitoring Committee
EASI	Eczema Area and Severity Index
EMR	Electronic Medical Records
GP	General practitioner
IQR	Inter-quartile range
ITT	Intention-to-treat
MCID	Minimum Clinically Important Difference
POEM	Patient Oriented Eczema Measure
RA	Research associate
SAP	Statistical Analysis Plan
SD	Standard Deviation
TCI	Topical calcineurin inhibitors
TSC	Topical corticosteroid cream

1. INTRODUCTION AND PURPOSE

This document details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from the BEE trial.

The purpose of the plan is to:

1. Ensure that the analysis is appropriate for the aims of the trial, reflects good statistical practice, and that interpretation of a priori and post hoc analyses respectively is appropriate.
2. Explain in detail how the data will be handled and analysed to enable others to perform the actual analysis in the event of sickness or other absence

Additional exploratory or auxiliary analyses of data not specified in the protocol are permitted but fall outside the scope of this analysis plan (although such analyses would be expected to follow Good Statistical Practice).

The analysis strategy will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will, if considered appropriate, be performed in accordance with the Analysis Plan, but if reported the source of such a post-hoc analysis will be declared.

Amendments to the statistical analysis plan will be described and justified in the final report of the trial.

2. SYNOPSIS OF STUDY DESIGN AND PROCEDURES

Eczema is a common disease in children and the majority have disease of mild or moderate severity which is diagnosed and managed exclusively in primary care (1). Clinical practice in this group of children is to prescribe a moisturiser (emollient) and topical corticosteroid cream (TSC)/topical calcineurin inhibitors (TCI) to use alongside to treat or prevent "flares" (2).

There are many different emollients available to buy over-the-counter and on prescription and a paucity of evidence that any one emollient is better than another. The main formulations are lotions, creams, gels and ointments, which vary in their consistency from "light" to "heavy". Research comparing the clinical effectiveness and acceptability of commonly used different emollients is therefore needed to provide evidence upon which clinicians and carers/patients can decide which emollient to try first.

The BEE study aims to compare the effectiveness and acceptability of four different types of emollient commonly used to treat eczema in children.

2.1 Trial objectives and aims

2.1.1 Primary objectives

The primary objective is to compare the medium term (over 16 weeks) effectiveness of the four main types of emollients in children with eczema with respect to patient-reported eczema symptoms.

2.1.2 Secondary objectives

The secondary objectives are to compare the emollient types with respect to:

- Patient-reported eczema symptoms (measured monthly over 52 weeks)
- Objective assessment of eczema signs by research nurse (at 16 weeks)
- Quality of life of the child (at 6,16 and 52 weeks)
- Impact of eczema on the family (at 16 and 52 weeks)
- Adverse events (over 16 and 52 weeks)
- Parent/carer satisfaction with study emollient (at 16 weeks)
- Frequency and quantity of study emollient and other emollient use (over 16 and 52 weeks)
- Use of other eczema treatments (including topical corticosteroids and topical calcineurin inhibitors) (over 16 and 52 weeks)
- Number of 'well-controlled' eczema weeks (over 16 and 52 weeks)

In a qualitative assessment – described in a separate qualitative analysis plan - the objectives are to:

- Understand and optimise recruitment processes
- Explore facilitators or barriers to study emollient use
- Explore carers' and children's experiences of study emollient use and their views about perceived effectiveness and/or acceptability of study emollients
- Contextualise the trial findings, as an aid to interpreting the results and their potential impact on clinical practice

2.2 Trial design

BEE is a pragmatic, multi-centre, individually randomised, parallel group superiority trial comparing four types of emollient in children with eczema, with an internal pilot and nested qualitative study.

2.3 Trial centres

BEE has three trial centres: i) Bristol, ii) Nottingham & Lincoln, and iii) Southampton.

2.4 Trial eligibility criteria

2.4.1 Trial inclusion criteria

Children eligible for inclusion in the study must:

- be aged between 6 months and less than 12 years
- have eczema diagnosed by an appropriately qualified healthcare professional (registered doctor, nurse or health visitor)
- have eczema rated mild or worse (Patient Oriented Eczema Measure (POEM) score >2) within 28 days of randomisation

The person giving consent for the child to participate in the study must:

- have parental responsibility for the participant
- be willing to use the randomly allocated emollient type as the only leave-on emollient for 16 weeks

2.4.2 Trial exclusion criteria

Children will be excluded from the study if:

- the child has a known sensitivity to any of the study emollients or their constituents
- the child is participating in another research study currently or in the last four months
- the child has any other known adverse medical or social circumstance that would make invitation to the study inappropriate (as determined by GP practice staff)

The child will be excluded from the study if the person giving consent:

- is unable to give informed consent
- has insufficient written English to complete outcome measures

2.5 Description of interventions

Participants will be randomised to one of four types of emollient:

1. Lotion
2. Cream
3. Gel
4. Ointment

Parents/carers are asked to agree to use their study emollient as the only leave-on emollient for 16 weeks. GPs are asked to prescribe them with directions to apply twice daily and when required, as per routine clinical practice. The amount of emollient prescribed during the study, by repeat prescriptions, will be determined by the family. Clinicians will be free to issue prescription for a smaller amount (e.g. 125g), if requested (e.g. for travel purposes).

If the family or their doctor/nurse judges that continuing their study emollient will be detrimental or the parent/child decides that they simply don't like it, they can stop using their allocated emollient and seek an alternative from their GP. In this instance, the GP/family will be encouraged to use another emollient that is of the same type. This will not affect their participation in the trial, and so they will continue to be followed up, unless they choose to withdraw at this or any other time.

Clinical management of eczema will otherwise be as usual – with treating clinicians and parents/carers free to make clinic appointments, referrals and to continue to use or change other treatments (including topical corticosteroids) as normal.

2.6 Randomisation procedures

Participants are randomised by members of the Bristol-based research team using a validated web-based randomisation system supplied by the Bristol Randomised Trials Collaboration (BRTC). Using computerised randomly-generated numbers, participants are randomised in a 1:1:1:1 ratio to the four groups, stratified by centre and minimised by baseline POEM (POEM score of 3-7 (mild) versus 17-28 (moderate to very severe)) and participant age (less than 2 years old versus 2 years and above).

2.7 Sample size and justification

As we have four groups, we conducted our sample size calculation to allow us sufficient power to pick up clinically meaningful differences in (n=6) pairwise comparisons after a global test by setting alpha to 0.05/6:

	Lotion	**Cream**	**Gel**	**Ointment**
Lotion				
Cream	Comparison 1			
Gel	Comparison 2	Comparison 3		
Ointment	Comparison 4	Comparison 5	Comparison 6	

Informed by prior feasibility study work (3), we aim to identify minimally clinically important difference (MCID) in POEM scores of 3.0 or more between any two treatment groups. Despite observing a SD of 4.89 in the feasibility trial (Choice of Moisturiser for Eczema Treatment (3), POEM score at 12 weeks among those whose self-referred and had a baseline POEM>2), we performed our sample size calculation using a standard deviation (SD) of 5.5 to allow for the observed SD to be greater than 4.89. This will also allow for smaller differences to be detected should the observed SD be less than 5.5. Based on these, we estimate we require 416 patients (104 in each group) to detect a difference of 3.0 in POEM scores between any two groups with 90% power and a significance level of 0.05 (after adjustment for multiple pairwise comparisons using the Bonferroni correction). This assumes equal numbers of children randomised to each group. To allow for 20% loss of follow-up, we propose recruiting 520 patients in total.

The 2012 paper by Schram and colleagues (4) determined a POEM MCID score of 3.4, but our POEM MCID of 3.0 is based on our feasibility trial data (5,6). In the feasibility study we employed five methods to determine POEM MCID (three anchor-based methods using the Parent Global Assessment as the anchor and two distribution-based methods), all suggesting a POEM score of around 3.0. While this is more conservative than the estimation by Schram et al, their data were from trials of adults with severe eczema, whereas ours used data from young children recruited in a primary care population – the majority of the participants were classified as suffering from moderate eczema (42%, baseline POEM score between 8-16). Designing the study to pick up a minimum difference of 3.0 will allow us to detect differences as small as this or larger differences as proposed by Schram et al.

2.8 Masking

The masking of all individuals involved in BEE is described below:

Participants (children and carers), treating clinicians, BRTC database staff, trial co-ordinator, trial administrator and qualitative research associate (RA)	**Not masked**
Supervisors of the qualitative RA	**Not masked** They will work with the qualitative RA to select participants for interviews and will study the interview transcripts (~60), but will not know the allocation of the remaining ~460 participants.
Clinical science officers (CSO)	**Masked** Masking will be monitored by self-report
Other Trial Management Group (TMG) members	**Masked**
Senior statistician	**Masked** The senior statistician – Stephanie MacNeill – has not seen any data when writing this SAP

	and will remain maskeded throughout the analysis only seeing aggregated data.
Trial Manager	**Not masked** The Trial Manager was masked prior to the writing of this SAP (v1) but after final approval will be unmasked to facilitate management of the study.
Trial statistician	**Not masked** The trial statistician was masked during the writing of this SAP (v1) knowing only an anonymised code for the different treatment groups. After final approval of the SAP, they will be unmasked to allow for detailed study of contamination and for the statistician to discuss unmasked data with the data monitoring committee as needed.

2.9 Trial committees

The study will be managed by a Trial Management Group (TMG) which will meet approximately monthly to monitor recruitment and data collection. The TMG will consist of all study investigators, the trial co-ordinator, research and administrative staff, with input from patient/public representatives.

Separate Trial Steering and Data Monitoring Committees (TSC and DMC) will meet regularly over the course of the study. The chief investigator (C-I) will attend all meetings, accompanied by the trial co-ordinator and other TMG/trial staff as appropriate.

The funder (National Institute for Health Research) will remotely monitor study progress against key targets by means of reports from the TMG and TS/DM-C.

2.10 Outcome measures

2.10.1 Primary outcome

The primary outcome is the POEM score measured weekly between weeks one and 16. POEM is a patient-reported outcome based on symptoms the previous week which is completed either by the child or their carer.

2.10.2 Secondary outcomes

The secondary outcomes for this study are:

- POEM measured monthly for 52 weeks
- Eczema signs (Eczema Area and Severity Index (EASI) by maskeded assessor) at 16 weeks
- Parent reported use of study emollient/other eczema treatments measured weekly for 16 weeks and monthly until 52 weeks
- Satisfaction with study emollient at 16 weeks
- Adverse effects collected weekly for 16 weeks and then monthly until 52 weeks
- Atopic Dermatitis Quality of Life (ADQoL) at 6, 16 and 52 weeks

- Child quality of life (Child Health Utility 9D, CHU9D) at 6,16 and 52 weeks
- Dermatitis Family Impact (DFI) at 16 and 52 weeks
- Proportion of weeks with well-controlled symptoms based on weekly POEM scores over the first 16 weeks

Data will also be collected on personal costs, healthcare contacts and prescriptions (by parent-report and review of participant's electronic medical records (EMR) after 52 weeks). Analysis of this data is dependent on securing other funding and, should we be successful, will be described in a separate health economics analysis plan.

2.11 Interim analysis

No interim analyses are planned.

3. GENERAL ANALYSIS CONSIDERATIONS

3.1 Analysis populations

The **Full Analysis set** includes all randomised participants. The primary intention-to-treat (ITT) analysis will be conducted using this dataset.

Per protocol analyses will be conducted on all participants in the Full Analysis set who are deemed to have adhered to their allocated emollient type during the first 16 weeks.

Safety analyses will be conducted on all randomised participants who receive at least one prescription for their allocated emollient.

3.2 Derived variables

The algorithms for the calculation of derived variables in this study are described below:

POEM	POEM is measured weekly and comprises seven questions relating to eczema symptoms over the past week. Each question carries equal weight and the responses to each question are scored from 0 to 4 as detailed below: ▪ 0 = no days ▪ 1 = 1-2 days ▪ 2 = 3-4 days ▪ 3 = 5-6 days ▪ 4 = Every day The overall POEM score is calculated as the sum across all seven questions, with a possible maximum of 28 points. (https://www.nottingham.ac.uk/research/groups/cebd/resources/poem.aspx)
EASI	The EASI score is calculated based on a physical assessment of the child's eczema by the study CSO. It incorporates both severity and extent of symptoms on different parts of the body. The head and neck, upper limbs, trunk and lower limbs are assessed separately for key signs of erythema (E; redness), induration/papulation/oedema (I; thickness), excoriation (Ex; scratching) and lichenification (L; lined skin) and rated on a scale of 0 (none) to 3 (severe) in steps of 0.5. Each sign is assessed for the entire body region and the percentage area affected within each body region is

	also assessed and scored (0: no active eczema; 1: 1-9%; 2: 10-29%; 3: 30-49%; 4: 50-69%; 5: 70-89%; 6: 90-100%). The EASI score for each body area is then calculated as: (E + I + Ex + L) x area category. The total EASI score is a weighted sum of the four EASI scores for each body area. The final EASI score ranges between 0 and 72. The weights allocated to each body area differ according to the age of the child. For children aged 7 years or under, head & neck area and upper extremities area are allocated weights of 0.2 each. The trunk area and lower extremities area are allocated weights of 0.3 each. For children aged 8 years and older, the head & neck area has a weight of 0.1, the upper extremities area has a weight of 0.2, the trunk area has a weight of 0.3 and the lower extremities area has a weight of 0.4 If a child turns 8 years during the trial we will use the same formula for the child throughout the study based on published guidance (http://www.homeforeczema.org/documents/easi-user-guide-jan-2017-v3.pdf)
ADQoL	The ADQoL is a preference-based quality of life measure which will be coded according to the developer's instructions (7)
CHU9D	The CHU9D is a nine-item preference-based measure of health-related quality of life measure which will be coded according to the developer's instructions (8)
DFI	The DFI is comprised of 10 questions designed to be completed with a one-week recall period. Each question is scored from 0-3 (0: not at all; 1: a little; 2: a lot; 3: very much). An overall score is obtained by summing across the 10 scores providing a range in values of 0-30.
Adherence	For each patient the number of days of self-reported use at least once of their allocated type of emollient is counted and expressed as the proportion of the total number of days for which non-missing emollient data are available.
Contamination	For each patient contamination is calculated as the proportion of days (among days where emollient data are available) where a non-allocated emollient type was used at least once.
Proportion of weeks with well-controlled symptoms	Between weeks 1-16, each week is classified as well-controlled or not based on whether the weekly POEM score is ≤2 (well-controlled) or >2 (not well controlled). The proportion of weeks with well-controlled symptoms is then calculated as the number of well controlled weeks divided by the number of weeks with non-missing POEM scores.

3.3 Procedures for missing data

3.3.1 Missing items in questionnaires

Missing questionnaire items will be summarised and the following rules will be used to calculate scores:

POEM	For missing items in the POEM questionnaire, we will adopt the following approach as recommended on the Centre for Evidence Based Dermatology website (https://www.nottingham.ac.uk/research/groups/cebd/resources/poem.aspx):

	• If one question is left unanswered this is scored 0 and the scores are summed and expressed as usual out of a maximum of 28 • If more than one question is left unanswered the questionnaire is not scored • If two or more response options are selected for a single question, the response option with the highest score should be recorded
DFI	If 1 or 2 items are missing, the missing items will be imputed by the participant mean for the remaining completed items. No imputation will be done if 3 or more items are missing.
ADQoL	No utility score will be calculated where items are missing
CHU9D	No utility score will be calculated where items are missing
Proportion of weeks with well-controlled symptoms	This variable is derived from weekly POEM scores and will only be missing if all POEM scores between weeks 1 and 16 are missing.

3.3.2 Missing baseline data

It is expected that baseline data will be largely complete for most variables and baseline POEM scores are required for randomisation so will be complete. For other outcome variables measured at baseline (EASI, ADQoL, DFI and CHU-9D) we will monitor missing baseline data and consider appropriate means of imputation in order to include these in our treatment efficacy models. We will follow the guidance of White and Thompson (10) in this matter.

3.3.3 Missing outcome data

We will explore patterns of missing in our primary outcome and consider possible mechanisms for this. Based on these and observed data, appropriate methods for imputing missing data will be considered in sensitivity analyses, including both "best" and "worst" case scenarios. Where assumptions are met, this may include multiple imputation by chained equations, for example. Should there be imbalance between treatment groups on important baseline characteristics, sensitivity analyses will be conducted where the main analysis will be repeated adjusting for these.

3.4 Study centre effects

Randomisation is stratified by centre and all multivariable analyses will adjust for centre.

3.5 Questionnaire completion

Outcomes are reported by patients weekly for the first 16 weeks then monthly. Efforts are made by the study team to collect questionnaires (online or on paper) until these are 2 weeks overdue.

3.6 Timing of final analysis

Final analysis of the primary outcome will be conducted after the last patient has reached the 16-week follow-up point.

4. DESCRIPTION OF PARTICIPANT CHARACTERISTICS

4.1 Disposition

A flow of patients through the trial will be summarised in a CONSORT diagram that will include the eligibility, reasons for exclusion, numbers randomised to the four treatment groups, losses to follow-up and the numbers analysed.

4.2 Baseline characteristics

Baseline characteristics of patients will be compared between the four arms by reporting relevant summary statistics in order to determine whether any potentially influential imbalance occurred by chance. Baseline characteristics will be summarised using means (SD), medians (Inter-quartile-range; IQR) or number (%) depending on the nature of the data and its respective distribution. If the baseline characteristics of any two groups differ by more than 10% or half a standard deviation then the effect of this variable on the outcome will be investigated in sensitivity analyses.

5. ASSESSMENT OF STUDY QUALITY

5.1 Eligibility checks

The numbers of patients and reasons for exclusion will be described.

5.2 Data management and data validation

The study follows the BRTC data management and quality management standard operating procedures (SOP-IT-004 and SOP-QM-001, respectively).

At the time of data entry, all data requested on the CRF will be recorded, checked and any missing data explained. A random sample of 10% of CRFs will be checked against the computerised database and relevant source data for quality purposes. This percentage will be increased if a significant error rate (more than 10% of those checked) is found.

Once the data is downloaded by the trial statistician internal consistency checks will be performed by them in preparing the data for analysis in Stata. They will aim to identify spurious values or inconsistencies in responses. When inconsistencies are identified, these will be reported to the trial manager who verifies the completed forms.

5.3 Adherence and contamination

The BEE internal pilot uses adherence and contamination as progression criteria. The definitions of both are outlined in section 3.2. For the purposes of progression, each participant will have a measure of contamination and adherence which is equal to the proportion of days where they were adherent (or used other emollients in the case of contamination). The medians of these measures will then be calculated for each treatment group.

We will use graphs to explore the patterns of study emollient use (ie. the study emollient was used on a given day), use of other leave-on moisturisers (ie. another leave-on study emollient of another type was used on a given day) and combinations of these. This will allow us to see how use changes over the duration of the study and may help us identify meaningful definitions of what constitutes “substantial contamination”. This could then inform further sensitivity analyses.

Should prescription record data be available at the end of the study we will also study the quantity of study emollient prescribed (grams/ml) and report this by treatment group using means (SD) or median (IQR) depending on the distribution of the data.

5.4 Protocol deviations

A protocol deviation is an unanticipated or unintentional divergence or departure from the expected conduct of a study inconsistent with the protocol, consent document or other study procedures. Of particular importance in the analysis are major deviations (violations) which may expose subjects to increased risk; compromise the integrity of the entire study or affect subject eligibility.

Protocol deviations will be described with information on the treatment group and nature of the deviation.

6. ANALYSIS OF EFFECTIVENESS

Stata version 15 (or higher) will be used for all BEE analyses. Two-tailed tests will be used with effect estimates, 95% confidence intervals (CI) and p-values presented. Analyses using regression models will adjust for stratification and minimisation variables as well as baseline values of the outcome studied. The primary approach for analysis will be on an intention-to-treat basis defined as analysing participants as randomised, regardless of the adherence to their allocated group. Mis-randomised patients will be analysed as randomised. Missing data will be imputed in sensitivity analyses.

6.1 Summary of primary and secondary outcomes

The primary and secondary outcomes assessed in BEE are described below:

Outcome	Measure	Timepoints	Interpretation	Range
Primary				
Patient-reported eczema symptoms	POEM	Baseline and weekly for 16 weeks	Numeric score where lower scores indicate fewer symptoms; higher scores more symptoms	0-28
Secondary				
Patient-reported eczema symptoms	POEM	Baseline and monthly for 52 weeks	*As above*	*As above*
Eczema severity (objective)	EASI	Baseline and at 16 weeks	Continuous score where lower scores indicate less severe eczema; higher scores more severe eczema	0-72
Adherence and contamination	-	Weekly for 16 weeks then monthly until 52 weeks	Treatment adherence and contamination are calculated as described in section 5.3	0-100
Parent-reported satisfaction with study emollient	-	16 weeks	Ordinal measure *	Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied
Adverse events	-	Weekly for 16 weeks then monthly until 52 weeks	Categorical measure	-

Disease-specific quality of life for the child	ADQoL	Baseline and at weeks 6, 16 and 52	Continuous utility score where a lower score indicates the worst health state(death); 1 being perfect health	0-1
Disease-specific quality of life for the family	DFI	Baseline and at weeks 16 and 52	Numeric score where a score of 0 indicates no impact on family life and 30 indicating maximum impact.	0-30
Generic quality of life for the child	CHU9D	Baseline and at weeks 6, 16 and 52	Continuous utility score where a lower score indicates the worst health state(death); 1 being perfect health	0-1
Proportion of weeks with well-controlled symptoms	-	Calculated from weekly POEM scores collected at weeks 1-16	Continuous	0-100

6.2 Primary analysis

The primary outcome is the weekly POEM score assessed for up to 16 weeks. For this analysis, a linear mixed model (weekly observations (level 1) nested within participants (level 2)) will be used to explore whether there are differences in mean POEM scores between treatment groups after adjusting for baseline scores and all stratification and minimisation variables used in the randomisation. This approach allows incomplete cases (ie. patients who did not complete all of their weekly scores) to contribute to the analysis. Therefore, all patients in the Full Analysis dataset who contributed at least one POEM score between weeks 1 to 16 will be included. The choice of covariance structure will be decided upon based on a review of the data.

We will check the normality assumptions of the residuals from the fixed part of the multi-level model and the random effects at the cluster level using graphs. Appropriate transformations will be considered if the assumptions of the model are not be met.

Pairwise comparisons will be conducted to identify which intervention groups differed. Graphs and tables will be used to present all six pairwise differences with confidence intervals. To account for multiple testing, a modified alpha of 0.0083 will be used for pairwise comparisons and 99.17% confidence intervals will be used to reflect this adjustment.

Note that during routine data verification it emerged that data entry errors had been made when baseline POEM categories were specified for the randomisation system. All regression models will include the stratified POEM scores as used in the randomisation; this includes primary, secondary and sensitivity analyses. Models of POEM scores additionally adjust for baseline POEM scores as a continuous variable and here the correct baseline POEM score will be used.

6.3 Sensitivity analyses of the primary outcome

A number of analyses are proposed to assess the sensitivity of the primary analysis to various assumptions. These are described below. Sensitivity analysis results will be represented alongside those of the primary analysis in order to be compared and contrasted. As sensitivity analyses will be exploratory in nature, 95% confidence intervals and p-values will be presented but will be interpreted with due caution.

6.3.1 Missing data

As described in section 3.3, we will conduct a sensitivity analysis imputing missing POEM scores for an analysis of the primary outcome.

6.3.2 Imbalance between treatment groups

Should there be evidence of imbalance between treatment groups on important baseline characteristics as described in section 4.2, sensitivity analyses will be conducted where the main analysis will be repeated adjusting for these.

6.3.3 Self-completed and CSO-collected POEM symptom data at 16-weeks

Between baseline and 16-weeks parents are invited to complete online or paper questionnaires regarding their child's eczema symptoms. At 16-weeks, the children are visited by the CSO where objective symptom severity data are collected (EASI). If the 16-week POEM data is not returned by the parent when this visit occurs, the CSO will collect POEM data for that week.

The primary analysis outlined in section 6.2 uses only the parent-completed symptom data. A sensitivity analysis will be conducted using the same modelling approach but incorporating the CSO-collected data at 16 weeks where parent-reported data are not available as well as an indicator variable for whether the data was parent-completed or CSO-collected.

Where a parent-completed measure at 16-weeks is ultimately returned and a CSO-collected is also provided these scores will be compared and differences described.

6.3.4 Per protocol analysis

As described in section 5.3, we are unable to pre-specify what constitutes "substantial contamination" prior to investigation of the data, but we will study patterns of emollient use over time to establish a meaningful definition. If contamination levels meet this definition, we will carry out a per protocol analysis.

6.3.5 Randomisation of ineligible participants

Over the course of data collected it emerged that a small number of randomised participants who engaged in follow-up were not in fact eligible due to data entry mistakes in their POEM scores at the screening stage. These participants will be included in the primary analysis, but a sensitivity analysis will be conducted where these participants will be excluded. The results of the two analyses will be compared.

6.4 Impact of COVID-19

On 11 March 2020, the WHO declared the COVID-19 outbreak as a global pandemic. A concern was raised by the study management group that increased hand washing and use of sanitiser gels might lead to worse eczema symptoms and reduce the effectiveness of the intervention. To explore the possible impact this had on symptoms, a sensitivity analysis will be conducted on the repeated measures POEM scores. This will first involve graphs summarising POEM scores before and during the period after the pandemic was declared. A binary variable classifying the follow-up period as pre-COVID19 and during will be generated. We will run a linear mixed model (weekly observations (level 1) nested within participants (level 2)) which will include a COVID-group interaction term to explore whether the differences in mean POEM scores between treatment groups differed between the period before and during COVID-19. The model will also account for baseline POEM scores and variables used in the randomisation.

6.5 Sub-group analyses

Pre-specified subgroup analyses will investigate whether treatment effectiveness (POEM), acceptability and quality of life are modified by the following factors measured at randomisation:

- Parent expectation: As the primary outcome is patient-reported and may be subject to performance bias, we will also explore whether reported effectiveness is linked to low or high expectation of effectiveness (pre-randomisation). Parents are asked to score on a scale of 1 ("very poor") to 5 ("very good") or "don't know" their thoughts on how effective they think different moisturisers are for treating the dry skin of eczema. For the purpose of sub-group analysis, the variable is classified as "poor" (score of 1 or 2), "average or unsure" (score of 3 or "don't know") or "good" (score of 4 or 5).
- Age: We would like to explore whether there are treatment differences in younger (<2 years) and older patients (≥2 years) (10).
- Disease severity: We would like to explore whether there are treatment differences between those with mild eczema versus those with moderate/severe eczema
- Diagnosis of eczema: We would like to explore whether there are treatment differences between children who do and do not fulfil the UK diagnostic criteria for atopic eczema

The statistical methods used in the primary analysis (multivariable regression with patient as a random effect)_will be extended to incorporate interaction terms with the treatment allocation, to test null hypotheses of no variation in treatment effect across subgroups.

6.6 Secondary outcomes

The approach for analysis of secondary outcomes will be on an intention-to-treat basis defined as analysing participants as randomised, regardless of the adherence to their allocated group. Secondary outcomes will be analysed according to the data type and frequency of recording. Continuous outcomes measured at multiple time points (POEM, ADQoL, DFI and CHU9D) will be analysed similarly to the primary outcome as described above.

The EASI score measured at baseline and 16 weeks will be analysed using a linear regression model adjusting for baseline values where available.

Parental satisfaction with the study emollient at 16 weeks will be analysed using an ordered logistic regression model adjusting for all stratification and minimisation variables. Results will be presented as odds ratios (ORs) and 95% confidence intervals.

Patterns of use of the study emollient and other eczema treatments – including topical corticosteroids - will be explored in the first instance using descriptive statistics. Based on these findings, comparisons may be made between treatment groups.

The proportion of weeks with well controlled symptoms will be analysed using a linear regression model adjusting for stratification and minimisation variables.

For each of the linear regression models used for secondary outcomes we will explore the assumptions of normally distributed residuals using graphs. Where assumptions are not met, transformations of the outcome will be considered. For mixed effect models, both the residuals of the fixed part of the model and the random effects will be studied.

Descriptive analysis of safety endpoints will be presented both according to randomised group and according to actual emollient used in the two groups.

6.6.1 Sensitivity analysis of the EASI score outcome

It was originally planned that – for each participant - eczema severity (EASI) would be assessed by the same CSO at baseline and at 16-weeks. This would address possible observer effects within children. Given recruitment patterns and staffing changes in local research teams, however, it is not always possible for the same person to collect these data at both time points.

We will report the number and proportion of children – by arm – who have had their 16-week and baseline EASI measured by the same CSO. The regression analysis of the outcome (outlined in section 6.5.) will also be replicated with additional adjustment for whether or not the same CSO collected the data at both time points.

6.7 Characteristics of non-study patients

Among those patients identified through GP record searches we will compare the age and gender of the following groups of children:

- those children deemed eligible for the trial compared to those deemed ineligible
- those children deemed eligible for the trial and agreed to further screening compared with those who declined further screening and those who did not respond

Of those agreeing to postal screening, we will compare the age and gender of those who were deemed eligible through this process with those deemed ineligible, those who declined and those who did not respond. Of those who completed postal screening, we will compare the age, gender and POEM scores of those who agreed to baseline assessments and those who declined.

We will also compare the age, gender and POEM score distributions between those children deemed eligible based on the screening questionnaire who consented to participate in the trial and those who declined to participate.

7. ANALYSIS OF SAFETY

All serious adverse events in the first 16 weeks and over the 52-week study period will be tabulated by allocated group. The number of events, number of patients having at least one event and the number of patients with more than one event will be tabulated. Serious adverse events will also be listed.

8. CHANGES MADE TO THE PLANNED STATISTICAL ANALYSES

All changes made to the planned statistical analyses are described below:

Previous version	Previous date	New version	New date	Brief summary of changes
1	9 August 2018	1.1	27 February 2020	Details of 2 sensitivity analyses are added in sections 6.3.3. and 6.5.1.
1.1	27 February 2020	1.2	16 October 2020	To ensure consistency with the protocol (v7) the term “blinding” was replaced with “masking” and we clarified where researchers were “not masked” rather than “partially blinded”

				The primary analysis was clarified in section 6.2 to reflect errors identified in the POEM scores used in the randomisation. A sensitivity analysis was described in section 6.3.5 to reflect how we propose to analyse participants randomised in error. A sensitivity analysis relating to the impact of COVID-19 was added to section 6.4 Time points in figure 2 were corrected

9. BIBLIOGRAPHY

1. Emerson RM, Williams HC, Allen BR. Severity distribution of atopic dermatitis in the community and its relationship to secondary referral. British Journal of Dermatology, 139, 73-6. *British Journal of Dermatology* 1998;139:73-76.

2. Health NCCfWsCs. CG057: Atopic eczema in children - management of atopic eczema in children from birth up to the age of 12 years. London: RCOG Press, 2007

3. Ridd MJ, Garfield K, Gaunt DM *et al*. Choice of Moisturiser for Eczema Treatment (COMET): feasibility study of a randomised controlled parallel group trial in children recruited from primary care. *BMJ Open* 2016; **6**(11). DOI: 10.1136/bmjopen-2016-012021

4. Schram ME, Spuls PI, Leeflang MM et al. EASI (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy* 2012; 67(1): 99-106.

5. Gaunt, DM, Metcalfe C and Ridd M. The Patient-Oriented Eczema Measure in young children: responsiveness and minimal clinically important difference. *Allergy* 2016; **71**: 1620-1625. Doi: 10.1111/all.12942

6. Howells L, Ratib S, Chalmers JR *et al*. How should minimally important change scores for the Patient-Oriented Eczema Measure be interpreted? A validation using varied methods. *British Journal of Dermatology* 2018; 178(5): 1135-1142.

7. Stevens KJ, Brazier JE, McKenna SP *et al*. The development of a preference-based measure of health in children with atopic dermatitis. *British Journal of Dermatology* 2005; 153(2): 372-7.

8. Stevens K. The Child Health Utility 9D (CHU9D) – A new Paediatric Preference Based Measure of Health Related Quality of Life. PRO Newsletter #43. May 2010.

9. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Statistics in medicine*. 2005 Apr 15;24(7):993-1007

10. Williams HC. Atopic dermatitis: The epidemiology, causes and prevention of atopic eczema. Cambridge: Cambridge University Press (2000).

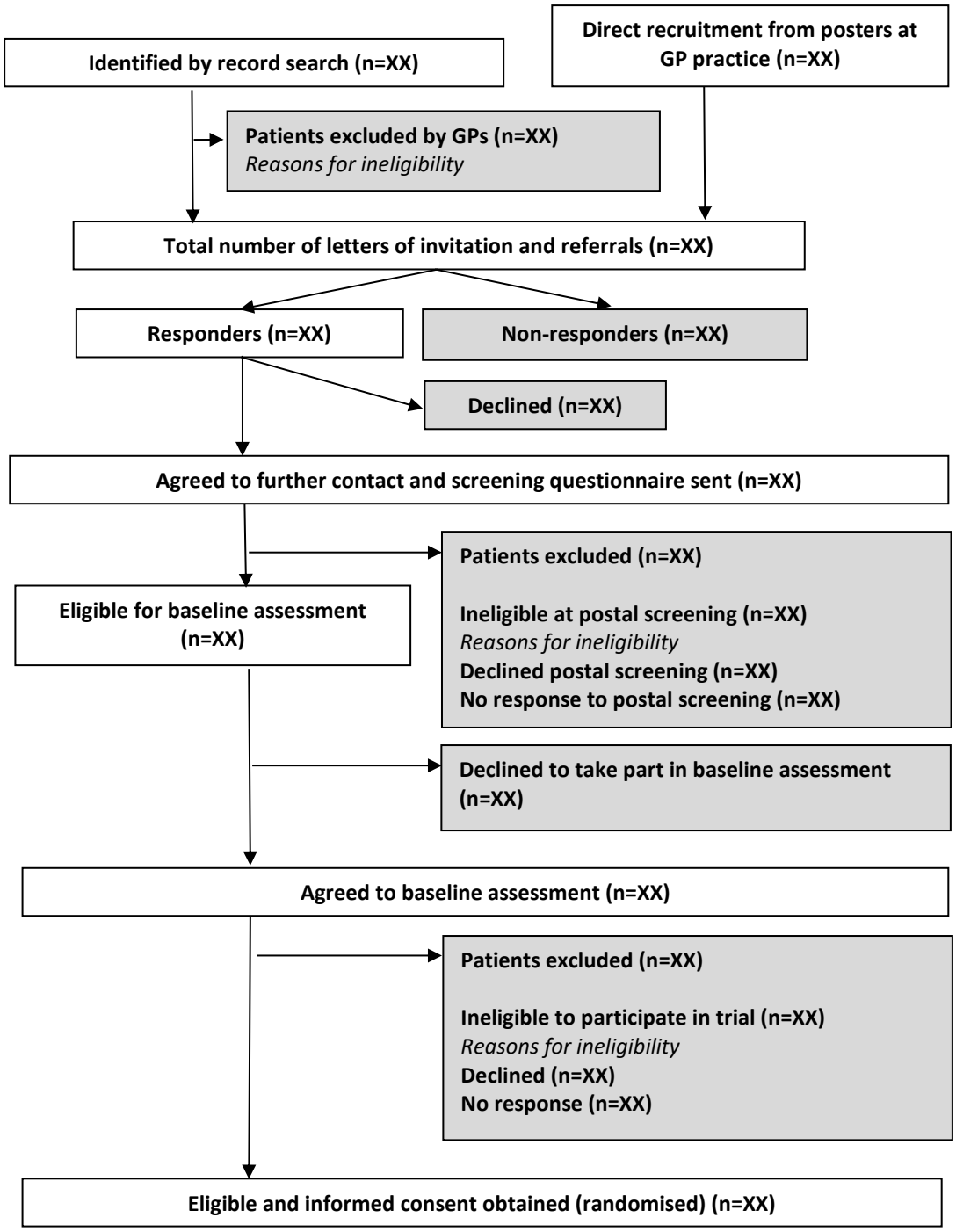
APPENDIX A: SKELETON TABLES AND FIGURES

Section	Outputs	
Section 1 Population	**Tables, figures and listings detailing the study population**	
	Figure 1	CONSORT diagram: recruitment pathway
	Figure 2	CONSORT diagram: randomisation onwards
	Table 1	Recruitment statistics by centre
	Table 2	Age and gender of children deemed eligible and those deemed ineligible at the GP records screening stage
	Table 3	Age and gender of children invited for further screening after being deemed eligible after the GP records screening stage
	Table 4	Age and gender of children who agreed to postal screening
	Table 5	Age, gender and POEM scores of children who completed postal screening
	Table 6	Age, gender and POEM scores at the screening stage of eligible children who consented to the study and those who did not
	Table 7	Protocol deviations
	Table 8	Details of individual protocol deviations
	Table 9	Withdrawal from the trial medication
	Table 10	Details of individual withdrawals from the trial medication
Section 2 Baseline data	**Summary tables of demographic information**	
	Table 11	**Baseline characteristics by treatment group**
Section 3 Outcomes	**Summary data and treatment estimates**	
	Table 12	Primary outcome: Mean weekly POEM scores over weeks 1-16
	Figure 3	Primary outcome: POEM scores over the 16-week primary outcome period by group
	Table 13	Completeness of individual items of the weekly POEM scores; n(%)
	Table 14	Primary outcome: Differences between treatment groups in weekly POEM scores over weeks 1-16
	Figure 4	Primary outcome: Adjusted differences (95% CI) between treatment groups in weekly POEM scores over weeks 1-16
	Table 15	Secondary outcome: Mean weekly POEM scores measured every four weeks over 52 weeks
	Figure 5	Secondary outcome: POEM scores over 52 weeks by group
	Table 16	Secondary outcome: Differences between treatment groups in POEM scores measured every four weeks over weeks 1-52
	Table 17	Secondary outcome: EASI at 16 weeks
	Table 18	Secondary outcomes: Description of ADQoL at each time point
	Table 19	Secondary outcome: DFI at 16 and 52 weeks
	Table 20	Secondary outcome: Satisfaction with emollient at 16 weeks
	Table 21	Secondary outcome: Proportion of weeks with well controlled symptoms
	Table 22	Sensitivity analysis: Differences between treatment groups in weekly POEM scores over weeks 1-16 additionally adjusting for variables showing imbalance at baseline
	Table 23	Number (%) of patients providing POEM scores by week

Section	Outputs	
	Table 24	Sensitivity analysis: Differences between treatment groups in weekly POEM scores over weeks 1-16 using imputed values where POEM scores are missing
Section 4 Safety data	**Summary tables and listings of all adverse events and serious adverse events**	
	Table 25	**Number of patients reporting adverse reactions by type and allocation**
	Table 26	**Number of patients reporting adverse reactions and number of events by type and treatment group**

A1. Population

Figure 1: CONSORT diagram: recruitment pathway



Notes:
Some patients may be ineligible for more than one reason

Figure 2: CONSORT diagram: randomisation onwards

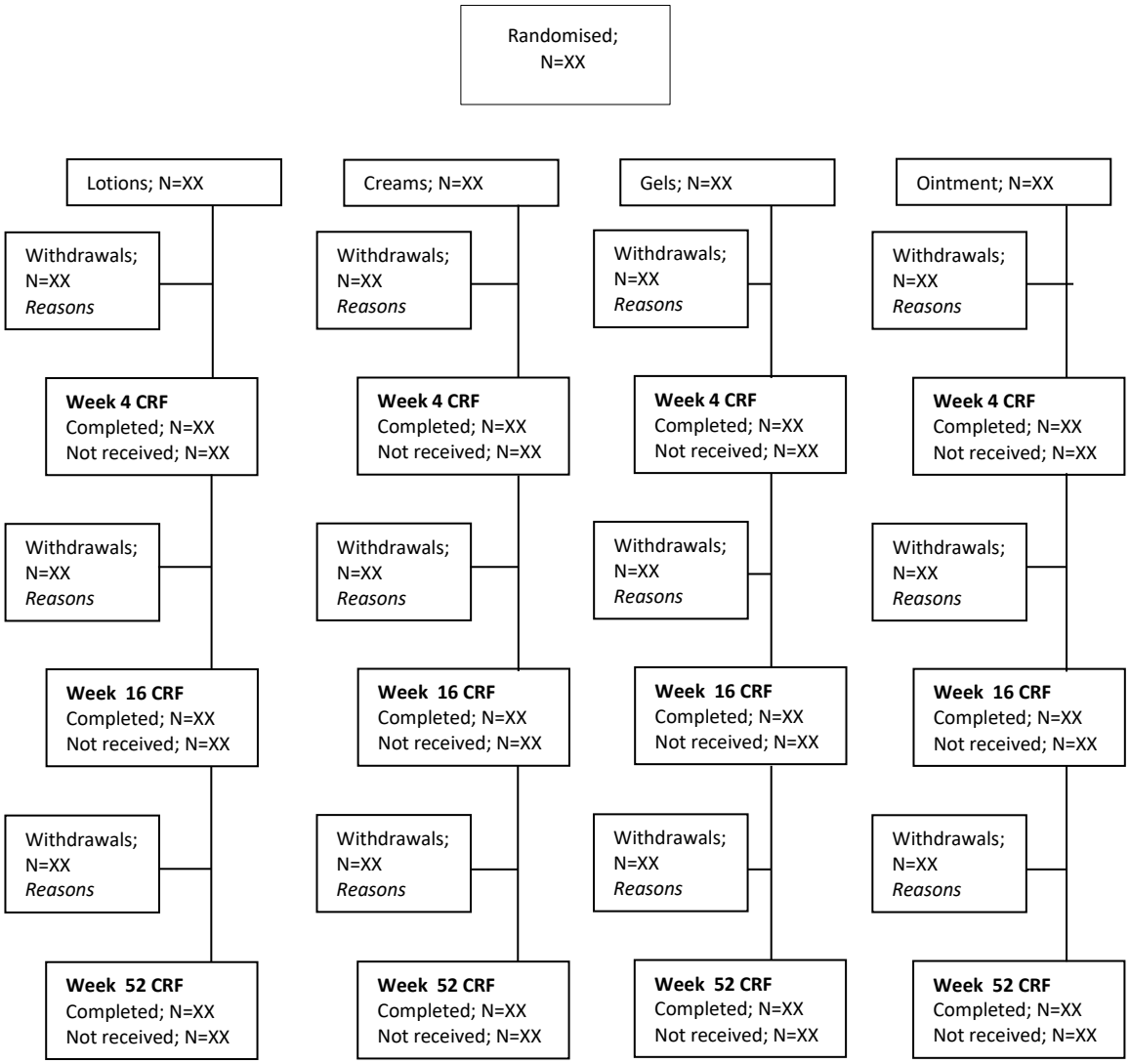


Table 1: Recruitment statistics by centre

	Bristol	Nottingham/Lincoln	Southampton	Total
Number of practices recruited				
Number of patients recruited				

Table 2: Age and gender of children deemed eligible and those deemed ineligible at the GP records screening stage

	Eligible (n=)	Not eligible (n=)
Age; mean (SD)		
Number female (%)		

Table 3: Age and gender of children invited for further screening after being deemed eligible after the GP records screening stage

	Responded to invitation (n=)		Did not respond (n=)
	Agreed to further screening (n=)	Declined further screening (n=)	
Age; mean (SD)			
Number female (%)			

Table 4: Age and gender of children who agreed to postal screening

	Eligible (n=)	Excluded (n=)		
		Ineligible (n=)	Declined (n=)	Did not respond (n=)
Age; mean (SD)				
Number female (%)				

Table 5. Age, gender and POEM scores of children who completed postal screening

	Agreed to baseline assessment (n=)	Declined to take part in baseline assessments (n=)
Age; mean (SD)		
Number female (%)		
POEM; mean (SD)		

Table 6: Age, gender and POEM scores at the screening stage of eligible children who consented to the study and those who did not

	Consented (n=)	Did not consent (n=)
Age; mean (SD)		
Number female (%)		
POEM; mean (SD)		

Table 7: Protocol deviations

	Randomised to lotion		Randomised to cream		Randomised to gel		Randomised to ointment		Overall	
	Patients	%	Patients	%	Patients	%	Patients	%	Patients	%
Any protocol deviation										

Nature of deviation										
..										

Table 8: Details of individual protocol deviations

Allocated treatment group	Centre	Further details (exact nature dependent on type of deviation)
…		

Table 9: Withdrawal from the trial medication

	Randomised to lotion		Randomised to cream		Randomised to gel		Randomised to ointment		Overall	
	Patients	**%**	**Patients**	**%**	**Patients**	**%**	**Patients**	**%**	**Patients**	**%**
Any withdrawal from the trial medication										
Nature of withdrawal										
Future questionnaires										
Child's note review for eczema and related consultation										
Use data already collected										
Reason given										
Study moisturiser not working/effective										
Adverse reaction to study moisturiser										
Disliked moisturiser given										
Changed mind										
Too many questionnaires										
Insufficient time										
Child's skin has improved										
Other										

Table 10: Details of individual withdrawals from the trial medication

Allocated treatment group	Days between randomisation and withdrawal from the trial medication	Patient withdrew consent	Reason	Completed further follow-up
…				

A2. Baseline data

Table 11. Baseline characteristics by treatment group

	Lotion (randomised n=)		Cream (randomised n=)		Gel (randomised n=)		Ointment (randomised n=)		Total (randomised n=)	
		N		N		N		N		N
Stratification variable: centre; n(%) Bristol Nottingham/Lincoln Southampton										
Minimisation variables										
Baseline POEM; n(%) Mild Moderate/severe										
Age; n(%) <2 years ≥2 years										
Demographic data										
Mean age (SD)										
Number female (%)										
Ethnic group; n(%) *White* *Black/African/Caribbean/Black British* *Asian/Asian British* *Mixed*										
Mean IMD score (SD)										
About the child's family										
Employment status of guardian completing questionnaire; n(%)										
Highest level of qualification or training of guardian completing questionnaire; n(%)										
Home ownership of guardian completing questionnaire; n(%)										
Eczema history										
Itchy skin the last year; n(%)										
Diagnosed food allergies; n(%) *No* *Yes* *Unsure/not diagnosed*										
Meeting diagnostic criteria; n(%)										
Prior eczema treatment										
Prior use of topical corticosteroids; n(%)										

Prior use of bath emollients; n(%)										
Opinions about study moisturisers										
Currently using lotions; n(%) *Never used* *Currently using* *Used previously* *Don't know*										
Currently using creams; n(%) *Never used* *Currently using* *Used previously* *Don't know*										
Currently using gels; n(%) *Never used* *Currently using* *Used previously* *Don't know*										
Currently using ointments; n(%) *Never used* *Currently using* *Used previously* *Don't know*										
Opinion on the effectiveness of lotions (1-5)*; mean (SD)										
Opinion on the effectiveness of creams (1-5)*; mean (SD)										
Opinion on the effectiveness of gels (1-5)*; mean (SD)										
Opinion on the effectiveness of ointments (1-5)*; mean (SD)										
Opinion on the acceptability of lotions (1-5)*; mean (SD)										
Opinion on the acceptability of creams (1-5)*; mean (SD)										
Opinion on the acceptability of gels (1-5)*; mean (SD)										
Opinion on the acceptability of ointments (1-5)*; mean (SD)										
Clinical data										
POEM score; mean (SD)										
POEM severity classification; n(%) Clear or almost clear Mild (3-7) Moderate (8-16) Severe (24-25) Very severe (26-28)										
DFI sore Mean (SD)										

Median (IQR)										
ADQoL; mean (SD)										
CHU-9D score; mean (SD)										
Mean Eczema Area and Severity Index (EASI) score (SD)										
EASI severity classification										
Number classified as clear or almost clear (%)										
Number classified as mild (%)										
Number classified as moderate (%)										
Number classified as severe or very severe (%)										

A3. Outcomes

Table 12 Primary outcome: Mean weekly POEM scores over weeks 1-16

	Allocated emollient							
	Lotion		Cream		Gel		Ointment	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Baseline								
Week 1								
Week 2								
Week 3								
Week 4								
Week 5								
Week 6								
Week 7								
Week 8								
Week 9								
Week 10								
Week 11								
Week 12								
Week 13								
Week 14								
Week 15								
Week 16								

Figure 3:Primary outcome: POEM scores over the 16-week primary outcome period by group
x-axis: week; y-axis: mean POEM score

Table 13 **Completeness of individual items of the weekly POEM scores; n(%)**

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Number with one missing item
Baseline Lotion Cream Gel Ointment								
Week 1 Lotion Cream Gel Ointment								
Week 2 Lotion Cream Gel Ointment								
Week 3 Lotion Cream Gel Ointment								

Week 4 Lotion Cream Gel Ointment								
Week 5 Lotion Cream Gel Ointment								
Week 6 Lotion Cream Gel Ointment								
Week 7 Lotion Cream Gel Ointment								
Week 8 Lotion Cream Gel Ointment								
Week 9 Lotion Cream Gel Ointment								
Week 10 Lotion Cream Gel Ointment								
Week 11 Lotion Cream Gel Ointment								
Week 12 Lotion Cream Gel Ointment								
Week 13 Lotion Cream Gel Ointment								
Week 14 Lotion Cream								

Gel Ointment								
Week 15 Lotion Cream Gel Ointment								
Week 16 Lotion Cream Gel Ointment								

Table 14 Primary outcome: Differences between treatment groups in weekly POEM scores over weeks 1-16

	Allocated emollient				**Univariate difference in mean POEM (95% CI)**	**Adjusted* difference in mean POEM (95% CI)**	**p-value**
	Lotion; mean (SD)	**Cream; mean (SD)**	**Gel; mean (SD)**	**Ointment; mean(SD)**			
POEM: 16 week repeated measures							(global test)
Pairwise comparisons							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

* Adjusted for baseline scores and all stratification and minimisation variables used in the randomisation

Figure 4: Primary outcome: Adjusted differences (95% CI) between treatment groups in weekly POEM scores over weeks 1-16
x-axis: comparison groups; y-axis: Difference in POEM score

Table 15: Secondary outcome: Mean weekly POEM scores measured every four weeks over 52 weeks

	Allocated emollient			
	Lotion; mean (SD)	**Cream; mean (SD)**	**Gel; mean (SD)**	**Ointment; mean(SD)**
Baseline				
Week 4				
Week 8				
Week 12				
Week 16				

Week 20				
Week 24				
Week 28				
Week 32				
Week 36				
Week 40				
Week 44				
Week 48				
Week 52				

Figure 5: Secondary outcome: POEM scores over 52 weeks by group
x-axis: week; y-axis: mean POEM score

Table 16 Secondary outcome: Differences between treatment groups in POEM scores measured every four weeks over weeks 1-52

	Lotion; mean (SD)	Cream; mean (SD)	Gel; mean (SD)	Ointment; mean(SD)	Univariate difference in mean POEM (95% CI)	Adjusted* difference in mean POEM (95% CI)	p-value
Secondary outcome – 52-week repeated measures							(global test)
Pairwise comparisons							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

* Adjusted for baseline scores and all stratification and minimisation variables used in the randomisation

Table 17: Secondary outcome: EASI at 16 weeks

	Lotion; mean (SD)	Cream; mean (SD)	Gel; mean (SD)	Ointment; mean(SD)	Univariate difference in mean EASI (95% CI)	Adjusted* difference in mean EASI (95% CI)	p-value
EASI score at baseline							
EASI score at 16 weeks							(global test)
Pairwise comparisons							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

* Adjusted for baseline scores and all stratification and minimisation variables used in the randomisation

Table 18: Secondary outcomes: Description of ADQoL at each time point

	Lotion		Cream		Gel		Ointment	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Baseline								
6 weeks								
16 weeks								
52 weeks								

Table 19: Secondary outcome: DFI at 16 and 52 weeks

	Lotion; mean (SD)	Cream; mean (SD)	Gel; mean (SD)	Ointment; mean(SD)	Univariate difference in mean DFI (95% CI)	Adjusted* difference in mean DFI (95% CI)	p-value
DFI score at baseline							
DFI score at 16 weeks							(global test)
DFI score at 52 weeks							(global test)
Pairwise comparisons: 16 weeks							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							
Pairwise comparisons: 52 weeks							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

* Adjusted for baseline scores and all stratification and minimisation variables used in the randomisation

Table 20: Secondary outcome: Satisfaction with emollient at 16 weeks

	Lotion; n(%)	Cream; n(%)	Gel; n(%)	Ointment; n(%)	Adjusted* difference (95% CI)	p-value
Satisfaction at 16 weeks Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						(global test)
Pairwise comparisons						
Lotion vs Cream Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						
Lotion vs Gel Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						
Lotion vs Ointment Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						
Cream vs Gel Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						
Cream vs Ointment Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						
Gel vs Ointment Very satisfied Mostly satisfied Neither satisfied/dissatisfied Dissatisfied Very dissatisfied						

* Adjusted for all stratification and minimisation variables used in the randomisation

Table 21: Secondary outcome: Proportion of weeks with well controlled symptoms

	Lotion; mean (SD)	Cream; mean (SD)	Gel; mean (SD)	Ointment; mean(SD)	Univariate difference in proportion of weeks with well controlled symptoms (95% CI)	Adjusted* difference in mean proportion of weeks with well controlled symptoms(95% CI)	p-value
Proportion of weeks with well controlled symptoms							(global test)
Pairwise comparisons							
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

* Adjusted for baseline scores and all stratification and minimisation variables used in the randomisation

Table 22: Sensitivity analysis: Differences between treatment groups in weekly POEM scores over weeks 1-16 additionally adjusting for variables showing imbalance at baseline

Comparison	Difference in mean POEM scores adjusting for baseline scores and all stratification and minimisation variables used in the randomisation (95% CI)	Difference in mean POEM scores adjusting for baseline scores and all stratification and minimisation variables used in the randomisation and any variables showing imbalance at baseline (95% CI)
Lotion vs Cream		
Lotion vs Gel		
Lotion vs Ointment		
Cream vs Gel		
Cream vs Ointment		
Gel vs Ointment		

Table 23: Number (%) of patients providing POEM scores by week

	Allocated emollient			
	Lotion; n(%)	**Cream; n(%)**	**Gel; n(%)**	**Ointment; n(%)**
Baseline				
Week 1				
Week 2				
Week 3				

Week 4				
Week 5				
Week 6				
Week 7				
Week 8				
Week 9				
Week 10				
Week 11				
Week 12				
Week 13				
Week 14				
Week 15				
Week 16				
Week 20				
Week 24				
Week 28				
Week 32				
Week 36				
Week 40				
Week 44				
Week 48				
Week 52				

Table 24: Sensitivity analysis: Differences between treatment groups in weekly POEM scores over weeks 1-16 using imputed values where POEM scores are missing

	Lotion; mean (SD)	Cream; mean (SD)	Gel; mean (SD)	Ointment; mean(SD)	Univariate difference in mean POEM (95% CI)	Adjusted* difference in mean POEM (95% CI)	p-value
Primary outcome – 16-week repeated measures Over the 16-week primary outcome period							(global test)
Lotion vs Cream							
Lotion vs Gel							
Lotion vs Ointment							
Cream vs Gel							
Cream vs Ointment							
Gel vs Ointment							

A4 Safety data

Table 25: Number of patients reporting adverse reactions by type and treatment group

	Allocated emollient			
	Lotion; n(%)	**Cream; n(%)**	**Gel; n(%)**	**Ointment; n(%)**
Any adverse events in the first 16 weeks				
Stinging *Itching* *Burning sensation* *Worsening of eczema* *Tingling* *Redness/inflammation* *Swelling* *Dryness* *Pain* *Peeling of the skin* *Skin infection* *Slip or fall*				
Any adverse events over the 52-week study period				

Stinging				
Itching				
Burning sensation				
Worsening of eczema				
Tingling				
Redness/inflammation				
Swelling				
Dryness				
Pain				
Peeling of the skin				
Skin infection				
Slip or fall				

Table 26: Number of patients reporting adverse reactions and number of events by type and treatment group

	Allocated emollient							
	Lotion; n		**Cream; n**		**Gel; n**		**Ointment; n**	
Events in the first 16 weeks	Number of events	Number of patients	Number of events	Number of patients	Number of events	Number of patients	Number of events	Number of patients
Stinging								
Itching								
Burning sensation								
Worsening of eczema								
Tingling								
Redness/inflammation								
Swelling								
Dryness								
Pain								
Peeling of the skin								
Skin infection								
Slip or fall								
Events over the 52-week study period								
Stinging								
Itching								
Burning sensation								
Worsening of eczema								
Tingling								
Redness/inflammation								
Swelling								
Dryness								
Pain								
Peeling of the skin								
Skin infection								
Slip or fall								