CONFIDENTIAL

STATISTICAL ANALYSIS PLAN FOR PROTOCOL 207619

A Clinical Study to Assess the Mildness of a Cosmetic cleanser in Healthy Subjects Using the Forearm-Controlled Application Technique (FCAT)

BIOSTATISTICS DEPARTMENT

GLAXOSMITHKLINE CONSUMER HEALTHCARE

Document type: Statistical Analysis Plan

Authors: PPD (Principal Statistician)

Document status: Final 1.0

Release date: 15-May-2017

The purpose of this Statistical Analysis Plan (SAP) is to describe the planned analyses and output to be included in the Clinical Study Report (CSP) for Protocol 207619. This SAP will be finalised prior to database freeze and treatment code un-blinding.

1 Study details

A cosmetic product that is freely available to the consumer must be safe when applied under normal or reasonably foreseeable conditions of use (ANVISA, 2012). Thus, the raw materials used in the product formulation must be of proven safety and with established use in the cosmetic industry. As a general requirement, the safety of the final formulation must also be confirmed before it is marketed.

Cleansers are designed to remove dirt, sweat, sebum and oils from the skin. This is achieved through the use of surfactants that aid in the uplifting of dirt and solubilisation of oils. However the interaction of cleansers and the stratum corneum can be detrimental to the skin, causing tightness and dryness as well as barrier damage, erythema, irritation and itch. (Wihelm, 1994)

Mildness is a factor that contributes to consumer acceptance of a personal cleansing product, especially in dry skin. The relative mildness (or irritation potential) of personal cleansing products is ideally judged under conditions of actual consumer use. However, these may not present the best conditions for discriminating product mildness differences, since normal use conditions typically do not induce differentiable product skin effects (Ertel, 1995).

The Forearm Controlled Application Technique (FCAT) is an exposure method that offers an efficient means to estimate the relative mildness (irritation) of personal cleansers through repeated overuse. It provides greater precision and sensitivity and minimises confounding effects due to biological diversity (Ertel, 1995). The FCAT is able to discriminate the relative mildness of personal cleansers in a wide variety of test conditions. Mildness can be defined by the parameters of skin redness and dryness (Lukacocvic, 1987).

Transepidermal water loss (TEWL) is the rate at which water permeates the stratum corneum and evaporates from the skin surface and qualifying TEWL can be used to support the examiner assessments.

Corneometry is a standard measurement of skin moisturisation which uses an electrical capacitance method to model the water content of the skin (Eisner et al, 1994). This assessment is included to provide an instrumental measurement of skin moisturisation to support the examiner assessments.

The objective of this clinical study is to assess the relative mildness of a cosmetic facial cleanser in comparison to water through repeated application to the volar forearm using the FCAT wash procedure. A positive control has been included in the design to help validate the trial results. Specifically, Imperial Leather Original bar soap has been selected as the positive control CCI [redacted] An area that will remain unwashed is also included in the study as a reference for the treated areas.

General safety and tolerability will be assessed based on the frequency and severity of Adverse Events (AEs).

1.1 Study design

Overall Design

This is a test site randomised, examiner blinded, positive and negative-controlled, single-center; Forearm Controlled Application Technique clinical study in healthy subjects to assess the mildness potential of a cosmetic facial cleansing product.

Subjects will give their written consent prior to any study procedures taking place. Visual assessments of dryness and redness will be conducted (by a trained blinded examiner). Each subject should have a score of zero for dryness and zero for redness on each volar forearm to be eligible for inclusion in the study. Eligible subjects will undergo a 5 to 7-day washout period, during which only the provided standard soap (Simple Soap®) will be used. This standard soap will also be used throughout the study in place of current cleansing product(s), but subjects should avoid using anything other than water on the volar forearms throughout the study, including the washout period.

A qualified dermatologist will assess subjects at Screening (Visit 1) for eligibility and again at Visit 2 for continued eligibility, to ensure the subjects are free of any clinically-relevant dermatological conditions. At the Baseline Visit (Visit 2, Day 1) examiner assessment of the dryness and redness of each proposed test site on each volar forearm will be performed prior to any wash application using the scoring system detailed in Appendix 2 of the Protocol. Each subject should have a score of zero for dryness and zero for redness to be considered eligible to continue in the study.

Prior to any wash procedure, subjects will be acclimatized to the environment of controlled temperature and humidity room and instrumental measurements of skin barrier function (measured with a Tewameter) and skin moisturisation (measured with a Corneometer) will be performed at Baseline (Visit 2 (AM)) in addition to the Baseline examiner ratings of dryness and redness at each forearm test site.

This will be followed by 5 days of repeated FCAT wash procedure where subjects will be instructed to attend the site for two wash visits per day for the first 4 days (Visits 2-5, AM and PM) and one wash visit on Day 5 (Visit 6, AM only). The AM and PM wash procedures on any given day will be separated by a minimum of 3 hours.

Visual assessment of dryness and redness will be conducted (by a trained, blinded examiner) prior to each PM wash procedure for Days 1-4 (Visits 2-5). The final examiner assessment of dryness and redness will be conducted on Day 5 (Visit 6), a minimum of 3 hours following the final Day 5 AM wash procedure.

A trained technician will perform each wash procedure to ensure consistency and eliminate the risk of cross-contamination of the study test sites. The technician will wash each of the

4 designated sites with one of the randomly allocated products; test product, positive control (Imperial Leather bar soap), negative control (sterile water only) or unwashed.

On Day 5, 3 hours after the last wash procedure, examiner ratings of dryness and redness and final instrumental measurements of skin barrier function and skin moisturisation will be performed. Additionally, a final clinical assessment by a qualified dermatologist will be performed to ensure that it is medically appropriate to discharge each subject from the study.

1.2 Study objectives

Objectives	Endpoints
Primary Objective	**Primary Endpoint**
• To compare change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5, for the test product versus water in the skin of healthy subjects.	• Examiner assessment of dryness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hours post last wash procedure.
Secondary Objectives	**Secondary Endpoint**
• To compare change from baseline in the examiner rating of redness to 3 hours following last wash procedure on day 5 for the test product versus water in the skin of healthy subjects.	• Examiner assessment of redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hours post last wash procedure.
• To compare change from baseline in the examiner rating of dryness and redness each day for test product and the positive control versus water in the skin of healthy subjects.	• Examiner assessment of dryness and redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 2, 3, 4 and 5 and 5, 3 hours post last wash procedure.
• To assess change from baseline in skin barrier function following the FCAT wash procedure for test product and positive control versus water.	• TEWL measurements on day 5, 3 hours post last wash procedure.
• To assess change from baseline skin moisturisation following the FCAT wash procedure for test product and positive control versus water.	• Corneometer measurements on day 5, 3 hours post last wash procedure.
• To evaluate the general safety of the test product.	• Assessment of frequency and severity of Adverse Events.

1.3 Treatments

	Test Product	Positive Control	Reference Product
Product Name	Micellar cleanser	Imperial Leather Original Bar Soap	Baxter Sterile Water

Product Formulation Code (MFC)	CCI [redacted]	N/A Market Place product	N/A Market Place product
Product Format	200 ml clear PET Bottle	100g Bar	1000ml Bottle
Application Quantity	0.09ml on to a moistened towel (with sterile water)	Moistened Towel(with sterile water) rubbed onto bar soap for 6 seconds to generate a lather	0.09ml on to a moistened towel (with sterile water)
Route of Administration	Topical dermal application		
Application Instruction	Applied on-site by technician		

Each subject will have four test sites; two test sites (3x3 cm) marked out on each volar forearm (Figure 1, Section 3.4). The distance between different application sites must be at least 1 cm to prevent either the test product, positive control (Imperial Leather bar soap) or negative control (sterile water) wash from spreading over and influencing neighbouring test sites. The site on each forearm for each wash application (or unwashed) will be determined according to the provided randomisation schedule. The randomisation schedule will identify the test sites for the test product, positive control and water (negative control) wash applications, and the unwashed site.

A trained technician will wash each marked site with the appropriate product.

All wash procedures will be performed at the study site by a technician to ensure consistent wash application and correct product usage. During the study, subjects will have the wash procedure completed four times daily at each allocated site, two wash applications in the morning (AM) and two wash applications in the afternoon (PM) on Days 1, 2, 3 and 4. On Day 5, subjects will have the wash procedure completed twice in the morning (AM) only. Each AM and PM wash procedure will be separated by a minimum of 3 hours.

1.4 Time points and visit windows

All data will be accepted for the analysis. Deviations from the scheduled assessment times are expected to be small and few. The following are the acceptable time windows.

Day#/Visit#	Activity	Time window
Day 1 (AM)/Visit 2	Wash procedures	At least 3 hours apart
Day 1 (PM)/Visit 2	Wash procedures	At least 3 hours apart
Day 2 (AM)/Visit 3	Wash procedures	At least 3 hours apart

Day#/Visit#	Activity	Time window
Day 2 (PM)/Visit 3	Wash procedures	At least 3 hours apart
Day 3 (AM)/Visit 4	Wash procedures	At least 3 hours apart
Day 3 (PM)/Visit 4	Wash procedures	At least 3 hours apart
Day 4 (AM)/Visit 5	Wash procedures	At least 3 hours apart
Day 4 (PM)/Visit 5	Wash procedures	At least 3 hours apart
Day 5 (AM)/Visit 6	Wash procedures	At least 3 hours apart

During screening subjects will sign an informed consent document and then a dermatological assessment will be conducted to ensure subjects have no dermatological conditions on their forearms that might impact subject safety or study results and to ensure subjects are classified as Fitzpatrick Phototype I to IV. A trained, blinded examiner will assess the dryness and redness of each volar forearm and only subjects with a score of zero for dryness and zero for redness will be enrolled into the study. Each subject's medical and medication history will be reviewed, in addition to the inclusion/exclusion criteria. Subsequently, site staff will review lifestyle guidelines and directions with eligible subjects.

A standard soap to be used in place of subjects' normal cleansers will be provided for use during the washout period and study treatment phase. Subjects will be instructed to avoid using any products, including the supplied cleanser, on their forearms throughout the washout and study treatment phase.

Consented eligible subjects will undergo a one-week washout period (5-7 days) prior to first wash procedure, during which the use of any skin cleansing products, including the standard soap, and any other topical products on the forearms will be prohibited.

On Day 1 (Visit 2, AM), eligible subjects will return to the study site and another dermatological assessment will be conducted, as will a review of any concomitant medications used since screening to ensure subject continued eligibility. Compliance concerning correct use of the standard soap will also be checked. Each subject will have two test sites (3x3 cm) marked out on each volar forearm. The distance between the two sites should be at least 1 cm to prevent either the test product, positive control product (Imperial Leather Original bar soap) or negative control (sterile water) applications from spreading over and influencing neighbouring test sites. A trained, blinded examiner assessment (Baseline; Day 1 - AM) of dryness and redness will be conducted at each marked site to ensure the dryness score is zero and the redness score is zero for each site for continued eligibility. The site on each forearm for each wash procedure (or unwashed) will be determined according to the provided randomisation schedule.

The randomisation schedule will determine the allocation of the 4 test sites for the test product, positive control and negative control wash applications and the site to be left unwashed. A trained technician will perform product application to ensure consistency and eliminate the risk of cross contamination of the study test sites.

Subjects will be acclimatised to the environment of a temperature (20-22ºC) and humidity (40-60% RH) controlled room for a period of at least 30 minutes before instrumental measurements of TEWL and corneometry at Baseline (Day 1 - AM) are taken, prior to any wash procedures and additionally at the final visit (Day 5 - PM) following the final wash procedure (Day 5 – AM). (EEMCO, 1997).

The treatment period will consist of 5 days of FCAT wash procedure where subjects will be instructed to come to the site for two wash procedures per day for the first 4 days (AM and PM) and one wash procedure on Day 5 (AM only). Each wash procedure will consist of two washings of each test site AM and PM. The AM and PM wash procedures on any given day will be separated by a minimum of 3 hours.

Further visual assessments of dryness and redness will be performed prior to each PM wash procedure on Days 1 to 4 and again at Day 5 (PM) following the final Day 5AM wash procedure (Appendix 2). The visual assessments of dryness and redness will be performed prior to subject acclimatisation, TEWL and Corneometry measurements on Day 1 and Day 5.

The FCAT generally produces only mild to moderate skin irritation, however, if any site reaches a dryness or redness score of ≥ 5.0 at any time during the study, treatment of all sites on that subject will be immediately discontinued, all test sites will be scored and these will be recorded as an Adverse Event and the subject will be withdrawn from the study.

2 Data analysis

Data analysis will be performed by Biostatistics Division in inVentiv International Pharma Services. Prior to database closure a Blind Data Review Meeting (BDRM) will be conducted in which various aspects of the trial will be discussed and agreed. The statistical analysis software used will be SAS Studio version 9.4 in a WINDOWS environment.

Except as described below, all listings will be produced for all randomised subjects.

2.1 Populations for analysis

Tables described in this section will be produced for all randomised subjects unless it is not mentioned otherwise in the specified sub-section.

2.1.1 Subject disposition

Subject disposition will be summarized as the number and percentage of subjects who complete the study, with the number who discontinue broken down by reason for discontinuation (Table 14.1.1) using all screened subjects.

2.1.2 Protocol Deviations

Important protocol violations (including violations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarised and listed.

Protocol violations will be tracked by the study team throughout the conduct of the study. Data will be reviewed prior to un-blinding and closure of the database to ensure all important violations are captured and categorised.

Major violations of the protocol procedures identified as liable to influence the efficacy outcome are as follows:
All-important violations will be defined in the "Review Listing Requirement (RLR)" document.

A list of protocol deviation will be listed in Listing 16.2.2.

2.1.3 Analysis populations

Four analysis populations are defined.

Population	Definition / Criteria	Analyses Evaluated
All Screened Subjects	• All subjects those who are screened	• Disposition
Randomised	• Randomised population will include all subjects who may or may not receive the any wash procedure.	• Protocol violation
Safety	• Safety population includes all subjects who received at least one wash procedure.	• Safety Analyses
Intent-to-Treat (ITT)	• The 'Intent to treat' (ITT) population includes all subjects who are randomised, receive at least one wash procedure, including sterile water, and have at least one post-baseline assessment available.	• Efficacy analysis (Dryness, Redness, etc.)

2.1.4 Subgroups/Stratifications

Not Applicable.

2.1.5 Centres pools

Not Applicable.

2.2 Patient demographics/other Baseline characteristics

Demographic and baseline characteristics summaries will be produced for the safety and ITT.

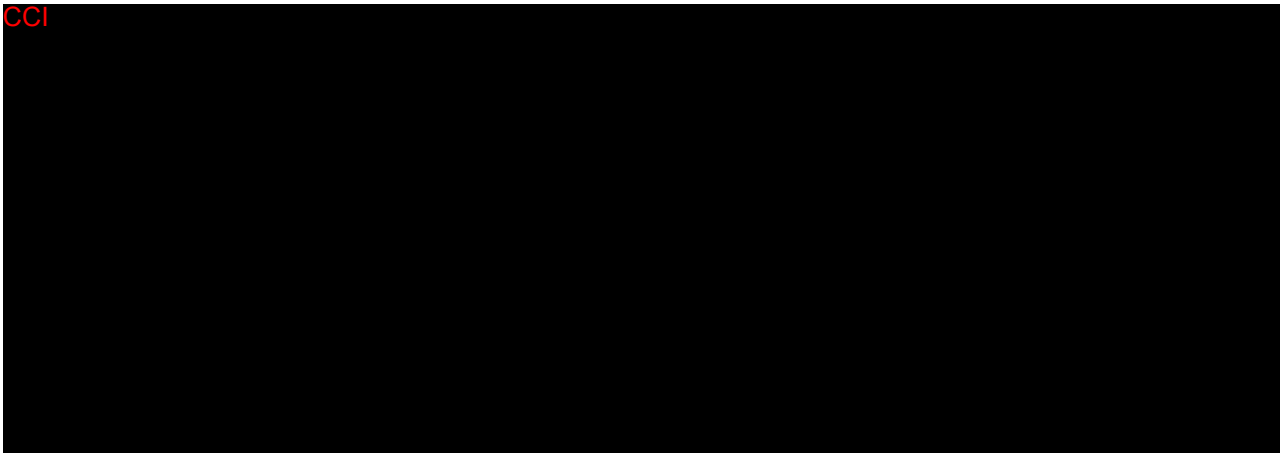
2.2.1 Demographic and Baseline characteristics

Categorical demographic variables include gender, race and Fitzpatrick scale. These variables will be summarized by the number and percentage of subjects with each relevant characteristic

(Table 14.1.4.1 and Table 14.1.4.2). Age will be summarized by the number of non-missing observations, mean, and standard deviation, median, minimum and maximum values.

The Fitzpatrick scale is a numerical classification that is widely used by dermatologists to classify a person's skin type by their response to the sun exposure (Fitzpatrick, 1988).

CCI



2.2.2 General medical history

Medical history data will not be presented in the study report. A data listing will be produced for evaluation of protocol violations only at the blinded data review stage.

2.2.3 Characteristics of Disease

Not Applicable.

2.3 Treatments (study treatment, other concomitant therapies, compliance)

Compliance data will be summarized for the ITT population. Exposure and other medications will be summarized on the safety population.

2.3.1 Study Product/treatment Compliance and Exposure

Not Applicable.

2.3.2 Any protocol deviation associated with bar soap wash adherence will be listed at the blinded data review stage. Concomitant medication

Concomitant medication/non-drug treatments data will not be presented in the study report. A data listing will be produced for evaluation of protocol violations only at the blinded data review stage.

2.4 Analysis of dermal response

2.4.1 Primary dermal response endpoint

2.4.1.1 Primary dermal response endpoint definition

The primary analysis will be performed to compare change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 for the test product versus water in the skin of healthy subjects. Analyses will be performed on the ITT population.

The forearm grading scale is described below –

Forearm Grading Scale:
Any skin response at any of the wash procedure areas on the volar forearm will be clinically assessed using the criteria recommended by Lukacovic, 1987. The following grades will be used to express the response observed at the time of examination:

CCI

CCI

Individual subjects observations for dryness and redness scores, TEWL and corneometer will be listed by visit and treatment (Listing 16.2.6).

Redness and dryness scores, TEWL and corneometer measurements at each time point and change from baseline in these scores and measurements at each visit will be summarized for test product, water, positive control and unwashed site using descriptive statistics such as the number of non-missing observations, mean, standard deviation, median, minimum, and maximum (Table 14.2.1.1, Table 14.2.1.2) using ITT population.

2.4.1.2 Statistical hypothesis, model, and method of analysis

The primary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of dryness.

The null and alternative hypotheses for the testing of non-inferiority will be following –

Null Hypothesis (H_{01}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water ≥ 0.25
Alternative Hypothesis (H_{a1}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water < 0.25

To test the above mentioned hypotheses an Analysis of Covariance (ANCOVA) model or appropriate statistical method will be used. The change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; baseline examiner rating of dryness as a covariate and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), and upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of dryness (Table 14.2.2.1).

2.4.1.3 Supportive analyses

To support parametric analysis approach described in the Section 2.4.1.2, a one-sided 0.05 level Wilcoxon sign rank test or other statistical method based will be performed on the adjusted differences. The adjusted differences for the test product will be calculated in the following way -

Adjusted differences of the test product at 3 hour on day 5 = Within subject change from baseline of 3 hour on day 5 – 0.25.

The result will be produced in the Table (Table 14.2.2.1).

Frequency and percentage of subjects with clinically relevant irritation responses will be provided by treatment. Clinically relevant irritation response for an individual subject is defined as, increase score in dryness from baseline of 1 or more with treatment (Table 14.2.2.2).

2.4.2 Secondary dermal response variables

The secondary efficacy variables are defined in the following sub-sections.

2.4.2.1 Secondary dermal response variable 1

The first secondary efficacy variable is examiner assessment of redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hour post last wash procedure.

2.4.2.2 Secondary dermal response variable 2

- The secondary efficacy variables are the examiner assessment of dryness and redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 2, 3 4 and 5, 3 hour post last wash procedure.
- Also a secondary efficacy variable is TEWL measurements on day 5, 3 hours post last wash procedure.
- Also a secondary efficacy variable is Corneometer measurements on day 5, 3 hours post last wash procedure.

2.4.3 Handling of missing values/censoring/discontinuations

Missing data will not be replaced or imputed. Dropouts will be included in analyses up to the point of discontinuation. Sensitivity analyses may be conducted if substantial number of subjects have discontinued from the study.

2.5 Analysis of secondary objectives

2.5.1 Dermal Response (secondary)

The first secondary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of redness.

To test the above mentioned hypotheses an Analysis of Covariance (ANCOVA) model will be used. The change from baseline in examiner rating of redness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; baseline examiner rating of redness as a covariate and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), upper one-sided 90% confidence limit will be presented. If the

upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of redness (Table 14.2.2.3).

Second, secondary endpoint examiner assessment of dryness (Table 14.2.2.4) and redness (Table 14.2.2.5)at day 2,3,4,and 5, 3 hour post last wash procedure will be analyzed using ANCOVA model for both dryness and redness separately. The model will be same as mentioned for the primary and first secondary efficacy endpoint analyses. Adjusted means, 95% CI and p-values will be provided for the following treatment comparison.

- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

Change from baseline in TEWL (Table 14.2.3.1) measurement and Corneometer measurement (Table 14.2.3.2) will be analysed using the ANCOVA model with treatment and site as fixed effect; baseline corresponding measurements as a covariate and subject as a random effect. The adjusted means, treatment differences, 95% CI and p-values will be presented for the following comparisons.

- Test product versus water
- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

If data are sufficiently non-normal in distribution, the non-parametric Wilcoxon signed ranks test or appropriate alternative will be used with median differences and 95% confidence interval presented based on the Hodges-Lehmann method.

No adjustment will be made for multiplicity testing for the secondary endpoints; however consistency of effects will be taken into consideration in interpreting the results.

2.5.2 Safety

2.5.2.1 Adverse events and Serious Adverse Events

All adverse events (AEs) will be listed in the (Listing 16.2.7.1) for the randomised subjects and all adverse events for the non-randomised subjects will be listed in Listing 16.2.7.2.

Treatment emergent adverse events (TEAEs) will be summarized by the number and percentage of subjects having any adverse event, an adverse event in each System Organ Class, and each individual adverse event (Table 14.3.1.1). All TEAEs will also be tabulated by severity and treatment (Table 14.3.1.2). Treatment-emergent AEs suspected of a relationship to study medication and those causing study discontinuation will be presented in a similar manner (Table 14.3.1.3). For treatment-related AEs, these will also be presented by severity and treatment, if applicable (Table 14.3.1.4).

Non-fatal serious adverse events and adverse events causing study treatment discontinuation will be listed (Listing 14.3.2.1).

2.6 Analysis of other variables

Not applicable.

2.7 Interim analysis

No interim analysis is planned.

2.8 Sample size calculation

The primary hypothesis is that the test product is not inferior to water in the change from baseline in the examiner rating of dryness on Day 5. Assuming that the standard deviation of the within subject treatment difference in change from baseline in examiner rating of dryness is 0.25 (based on previous study GSKCH: CCI), CCI [redacted]

A sufficient number of subjects will be screened in order to randomise 45 subjects and to ensure that at least 40 subjects complete the study.

3 Changes to the Protocol Defined Statistical Analysis Plan

There were no changes or deviations to the originally planned statistical analysis specified in the protocol version 1.0 [(Dated: 06/MAR/2017)].

4 Appendix 1:

4.1 List of Tables, Listings and Figures

4.2 Tables

Table Number	Table Title (Population)	Template
14.1.1	Subject Disposition (All Screened Subjects)	Appendix 2
14.1.4.1	Subject Demographics and Baseline Characteristics (Safety Population)	Appendix 2
14.1.4.2	Subject Demographics and Baseline Characteristics (ITT Population)	14.1.4.1
14.2.1.1	Summary of Examiner assessment rating on dryness and redness by Visit and Treatment (ITT Population)	Appendix 2
14.2.1.2	Summary of TEWL and Corneometer measurements by Visit and treatment (ITT Population)	Appendix 2
14.2.2.1	Statistical Analysis of examiner assessment of dryness on day 5, 3 hour post last wash procedure (ITT Population)	Appendix 2
14.2.2.2	Summary of clinically relevant examiner assessment of dryness and redness on day 2,3,4,and 5, 3 hour post last wash procedure by treatment (ITT Population)	Appendix 2
14.2.2.3	Statistical Analysis of examiner assessment of redness on day 5, 3 hour post last wash procedure (ITT Population)	14.2.2.1
14.2.2.4	Statistical Analysis of examiner assessment of dryness on day 2,3,4 and 5, 3 hour post last wash procedure (ITT Population)	Appendix 2
14.2.2.5	Statistical Analysis of examiner assessment of redness on day 2,3,4 and 5, 3 hour post last wash procedure (ITT Population)	14.2.2.1
14.2.3.1	Statistical Analysis of TEWL measurement on day 5, 3 hour post last wash procedure (ITT Population)	14.2.2.1
14.2.3.2	Statistical Analysis of Corneometer measurement on day 5, 3 hour post last wash procedure (ITT Population)	14.2.2.1
14.3.1.1	Treatment emergent Adverse Event (Safety Population)	Appendix 2
14.3.1.2	Treatment emergent Adverse Event by Severity and Treatment (Safety Population)	Appendix 2

Table Number	Table Title (Population)	Template
14.3.1.3	Treatment emergent Treatment Related Adverse Event (Safety Population)	14.3.1.1
14.3.1.4	Treatment emergent Treatment Related Adverse Event by Severity and Treatment (Safety Population)	14.3.1.2

4.3 Listings

Listing Number	Listing Title (Population)	Template
14.3.2.1	Listing of Serious Adverse Events leading to Discontinuation (Randomised population)	16.2.7.1
16.1.7.	Randomization information (Randomised Population)	Appendix 2
16.1.9.1	Statistical Analysis of examiner assessment of dryness on day 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.1.9.2	Statistical Analysis of examiner assessment of redness on day 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.1.9.3	Statistical Analysis of examiner assessment of dryness on day 2,3,4 and 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.1.9.4	Statistical Analysis of examiner assessment of redness on day 2,3,4 and 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.1.9.5	Statistical Analysis of TEWL measurement on day 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.1.9.6	Statistical Analysis of Corneometer measurement on day 5, 3 hour post last wash procedure (ITT Population)	SAS output for model analysis
16.2.2	Individual Subjects Protocol Deviations (Randomised Population)	Appendix 2
16.2.6	Subjects with Skin dryness, redness rating and measurement on TEWL and corneometer (ITT Population)	Appendix 2
16.2.7.1	Listing of All Adverse Events (Randomised Population)	Appendix 2
16.2.7.2	Listing of All Adverse Events (Non-randomised Subjects)	16.2.7.1

Note: If there are no data to display generate a null listing.

4.4 Top line Outputs:

Table/Listing Figure Number	Table/Listing/Figure Title (Population)
14.1.1.1	Subject Disposition (All Screened Subjects)
14.1.4.2	Subject Demographics and baseline Characteristics (ITT Population)
14.2.2.1	Statistical Analysis of examiner assessment of dryness on day 5, 3 hour post last wash procedure (ITT Population)
14.2.2.2	Statistical Analysis of examiner assessment of redness on day 5, 3 hour post last wash procedure (ITT Population)
14.3.1.1	Treatment emergent Adverse Event (Safety Population)
16.2.7.1	Listing of All Adverse Events (Randomised Population)

5 Appendix 2:

5.1 Templates for the Tables, listings and figures

This is a guideline which will give the guidance of treatment labels that will be used for the table header and in the figures, listings and in the footnotes.

The treatment labels for the column heading will be as follow:

- Micellar cleanser
- Imperial Leather Original Bar Soap
- Baxter Sterile Water
- Unwashed

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.1.1
Subject Disposition
All Screened Subjects

All Screened Subjects (N=XX)

	Overall(N=XX) n (%)
TOTAL NUMBER OF SUBJECTS SCREENED	xx (xx.x)
SUBJECTS NOT RANDOMISED	xx (xx.x)
DID NOT MEET STUDY CRITERIA	xx (xx.x)
ADVERSE EVENTS	xx (xx.x)
ETC.	xx (xx.x)
SUBJECTS RANDOMISED	xx (xx.x)
COMPLETED	xx (xx.x)
DID NOT COMPLETE	xx (xx.x)
ADVERSE EVENT	xx (xx.x)
LOST TO FOLLOW UP	xx (xx.x)
PROTOCOL DEVIATION	xx (xx.x)
WITHDRAWAL OF CONSENT	xx (xx.x)
OTHER	xx (xx.x)
RANDOMISED POPULATION	xx (xx.x)
SAFETY POPULATION	xx (xx.x)
INTENT TO TREAT POPULATION	xx (xx.x)

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.1.4.1
Subject Demographics and Baseline Characteristics
Safety Population

Safety Population (N=XX)

Demographic variables	Overall (N=XX)
SEX n (%)	
MALE	xx (xx.x)
FEMALE	xx (xx.x)
RACE n (%)	
ASIAN	
BLACK OR AFRICAN AMERICAN	
AMERICAN INDIAN OR ALASKA NATIVE	xx (xx.x)
NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	xx (xx.x)
WHITE	xx (xx.x)
MULTIPLE	xx (xx.x)
AGE (YEARS)	
N	XX
MEAN	XX.X
SD	XX.XX
MEDIAN	XX.X
MINIMUM	XX
MAXIMUM	XX
FITZPATRICK SCALE FOR SKIN TYPE	
I	xx (xx.x)

II	xx (xx.x)
III	xx (xx.x)
IV	xx (xx.x)
V	xx (xx.x)
VI	xx (xx.x)

Fitzpatrick Scale For Skin Type: I= Always burns easily; never tans (pale white skin);II= Always burns easily; tans minimally (white skin);III= Burns moderately; tans gradually (light brown skin); IV= Burns minimally, always tans well (moderate brown skin);V= Rarely burns, tans profusely (dark brown skin); VI= Never burns (deeply pigmented dark brown to black skin)

PPD

Programming Note:

- Repeat Table 14.1.4.2 with ITT Population.
- Also, for the corneometer assessment

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.1.1
Summary of Examiner Assessment rating on Dryness, Redness by Visit and Treatment
Intent to Treat Population

Intent to Treat Population N=XX)

Day/Visit/Assessment		Micellar cleanser (N=XX)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
		Raw	Change	Raw	Change	Raw	Change	Raw	Change
Day 1 (AM)/Visit 2/Dryness	N	xx		xx		xx		xx	
	MISSING	xx		xx		xx		xx	
	SCORE=0	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=1	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=2	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=3	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=4	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=5	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=6	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	MEAN	xx.xx		xx.xx		xx.xx		xx.xx	
	SD	xx.xxx		xx.xxx		xx.xxx		xx.xxx	
Day 1 (PM)/Visit 2/Dryness	Same as above								
Day 2 (PM)/Visit 3/Dryness	Same as above								
.....									
Day 5 (PM)/Visit 6/Dryness	Same as above								
Day 1 (AM)/Visit 2/Redness	Same as above								
.....									
Day 5 (PM)/Visit 6/Redness	Same as above								

Dryness rating: 0= No dryness; 1.0=Patches of slight powederiness and occasional patches of small scales may be seen; 2.0= Generalised slight powederiness; minimal popular response; 3.0= Generalised moderate powederiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powederiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness; 6.0=Severe redness.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.1.2
Summary of TEWL (g/m^2/hr) and Corneometer (i.u.) measurements by Visit and Treatment
Intent to Treat Population

Intent to Treat Population N=XX)

Day (AM/PM)/Visit/Assessment	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	Raw	Change	Raw	Change	Raw	Change	Raw	Change
Day 1 (AM)/Visit 2/TEWL								
n	xx		xx		xx		xx	
Missing	xx		xx		xx		xx	
Mean	xx.x		xx.x		xx.x		xx.x	
SD	xx.xx		xx.xx		xx.xx		xx.xx	
SE	xx.xx		xx.xx		xx.xx		xx.xx	
Median	xx.x		xx.x		xx.x		xx.x	
Minimum	xx		xx		xx		xx	
Maximum	xx		xx		xx		xx	

PPD

Programming Note:

- Repeat for Day 5
- Also, for the corneometer assessment

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.2.1
Statistical Analysis of examiner assessment of dryness on day 5, 3 hour post last wash procedure
Intent to Treat Population

Intent to Treat Population N=XX)

	Micellar cleanser (N = xx)		Baxter Sterile Water (N = xx)	
	Baseline [1]	Change	Baseline [1]	Change
Day 5, 3 hour post last wash				
n	xx	xx	xx	xx
Missing	xx	xx	xx	xx
Mean	xx.x	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx	xx.xx
SE	xx.xx	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x	xx.x
Minimum	xx	xx	xx	xx
Maximum	xx	xx	xx	xx
Adjusted Mean (SE) [2]		xx.x (xx.xx)		xx.x (xx.xx)
95% CI [2]		(xx.x, xx.x)		(xx.x, xx.x)
P-value [2]		0.xxxx		0.xxxx

Comparison Between Treatments	Difference [2,3]	95% CI	P-value
Micellar cleanser vs Baxter Sterile Water	xx.x	(xx.x, xx.x)	0.xxxx
Supportive Non-Parametric Analysis [4]			0.xxxx

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.

[2] From ANCOVA model with change from baseline in examiner assessment on dryness as response and treatment, site as fixed effects and examiner assessment on dryness as covariates. If one-sided upper 95% CI is less than 0.25 then non-inferiority is established.

[3] Difference is first named dentifrice minus second named dentifrice such that a negative difference favours first named dentifrice.

[4] P-value from Wilcoxon sign rank test, which is based on the adjusted differences. Adjusted difference is calculated in the following manner;

Adjusted differences of the test product at 3 hour on day 5 = Change from baseline of 3 hour on day 5 – 0.25.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.2.2
Summary of clinically relevant examiner assessment of dryness and redness on day 2,3,4,and 5, 3 hour post last wash procedure by treatment
Intent to Treat Population

Intent to Treat Population (N=XX)
Treatment:Miscellar Cleanser

Day # (time)/Visit #	Rating	Dryness Rating ----------Baseline-------							Redness Rating ----------Baseline-------						
		0	1.0	2.0	3.0	4.0	5.0	6.0	0	1.0	2.0	3.0	4.0	5.0	6.0
Day 1(PM)/Visit 2	0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	1.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	2.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	3.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	4.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	5.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	6.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Day 2 (PM)/Visit 3															
.....															
Day 5 (PM)/Visit 6															

Dryness rating: 0= No dryness; 1.0=Patches of slight powederiness and occasional patches of small scales may be seen; 2.0= Generalised slight powederiness; minimal popular response; 3.0= Generalised moderate powederiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powederiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness;6=Severe redness.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.2.4
Statistical Analysis of examiner assessment of dryness on day 2, 3, 4 and 5, 3 hour post last wash procedure
Intent to Treat Population

Intent to Treat Population N=XX)

	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	Baseline [1]	Change	Baseline [1]	Change	Baseline [1]	Change	Baseline [1]	Change
Day 2, 3 hour post last wash								
n	xx	xx	xx	xx	xx	xx	xx	xx
Missing	xx	xx	xx	xx	xx	xx	xx	xx
Mean	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
SE	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
Minimum	xx	xx	xx	xx	xx	xx	xx	xx
Maximum	xx	xx	xx	xx	xx	xx	xx	xx
Adjusted Mean (SE) [2]		xx.x (xx.xx)		xx.x (xx.xx)	xx.x (xx.xx)		xx.x (xx.xx)	
95% CI [2]		(xx.x, xx.x)		(xx.x, xx.x)	(xx.x, xx.x)		(xx.x, xx.x)	
P-value [2]		0.xxxx		0.xxxx	0.xxxx		0.xxxx	

Comparison Between Treatments	Difference [2,3]	95% CI	P-value
Micellar cleanser vs Baxter Sterile Water	xx.x	(xx.x, xx.x)	0.xxxx
Imperial leather Original bar Soap vs. Baxter Sterile Water			
Imperial leather Original bar Soap vs. Micellar cleanser			
Baxter Sterile Water vs.Un-washed			
Supportive Non-Parametric Analysis [4]		(xx.x, xx.x)	0.xxxx

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.

[2] From ANCOVA model with change from baseline in examiner assessment on dryness as response and treatment, site as fixed effects and examiner assessment on dryness as covariates. If one-sided upper 95% CI is less than 0.25 then non-inferiority is established.

[3] Difference is first named dentifrice minus second named dentifrice such that a negative difference favours first named dentifrice.

[4] P-value from Wilcoxon sign rank test, 95% CI is constructed on the Hodges-Lehmann method.

PPD

Programming Note:

- Repeat for day 3, 4 and 5, 3 hour post last wash
- Non-parametric method will be presented only if there is violation of normality.
- For the Table 14.2.2.5 the footnotes will be modified as following –

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.

[2] From ANCOVA model with change from baseline in examiner assessment on redness as response and treatment, site as fixed effects and examiner assessment on redness as covariates.

[3] Difference is first named dentifrice minus second named dentifrice such that a negative difference favours first named dentifrice.

[4] P-value from Wilcoxon sign rank test, 95% CI is constructed on the Hodges-Lehmann method.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.3.1.1
Summary of Treatment emergent Adverse Event
Safety Population

Safety Population: N=xx

System Organ Class Preferred Term	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE
NUMBER OF SUBJECTS WITH AT LEAST ONE AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NUMBER OF SUBJECTS WITH NO AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN RELATED AES	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN AND SUBCUTANEOUS TISSUE DISORDERS	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
ERYTHEMA	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
DERMATITIS	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NON SKIN RELATED AES	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
GASTROINTESTINAL SYSTEM	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
ABDOMINAL PAIN	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
DRY MOUTH	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
VOMITTING	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx

n (%) = Number (percent) of subjects nAE = Number of adverse events.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.3.1.2
Summary of Treatment emergent Adverse Event by Severity and Treatment/Test Site
Safety Population

Safety Population: N=xx

System Organ Class Preferred Term	Micellar cleanser (N = xx)						Imperial Leather Original Bar Soap (N = xx)						Baxter Sterile Water (N = xx)						Unwashed (N=xx)					
	Mild		Moderate		Severe		Mild		Moderate		Severe		Mild		Moderate		Severe		Mild		Moderate		Severe	
	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE
NUMBER OF SUBJECTS WITH AT LEAST ONE AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NUMBER OF SUBJECTS WITH NO AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN RELATED AES																								
SKIN AND SUBCUTANEOUS TISSUE DISORDERS																								
ERYTHEMA																								
NON SKIN RELATED AES																								
GASTROINTESTINAL SYSTEM																								
ABDOMINAL PAIN																								

n (%) = Number (percent) of subjects nAE = Number of adverse events.

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.1.7
Randomisation information
Randomised Population

Subject Number	Age/Sex/Race[1]	Randomization Number	Date of Randomisation	Treatment Randomised
PPD				

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

PPD

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.2
Individual Subjects Protocol Deviations
Randomised Population

Subject Number	Age/Sex/Race[1]	Visit #	Deviation Sequence	Protocol Deviation
PPD				xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

PPD

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.6
Subjects with Dermal Response and Superficial irritation Responses
ITT Population

Subject Number	Age/Sex/Race[1]	Phase/Visit #	Treatment	Dryness Rating[2]	Redness Rating[3]	TEWL measurement ($g/m^2/hr$)[4]	Corneometer (i.u.)[5]
PPD							xxxxxxxxxxxxxxxxxxxxx

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

[2] Dryness rating: 0= No dryness; 1.0=Patches of slight powederiness and occasional patches of small scales may be seen; 2.0= Generalised slight powederiness; minimal popular response; 3.0= Generalised moderate powederiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powederiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

[3]Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness; 6=Severe redness.

[4] g=gram; m^2=square meter; hr=hour

[5] i.u.= instrumental units

PPD

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.7
Listing of All Adverse Event
Randomised Population

Treatment Group: Micellar cleanser

Subject Number	Age/Sex/Race [1]	Site	Adverse Event (Preferred Term) (System Organ Class)	Start Date /Study Day[2]	Start Time	End Date	End Time	Frequency /Intensity[3]	Related to Study Product?	Action Taken re Study Product	Outcome	Serious?	Withdrew?[4]
PPD													

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.
[2] Study day is the day relative to start of treatment, day 1 being the day of first treatment.
[3] INT = Intermittent and SGLE = Single.

[4] Did subject withdraw from study as a result of this adverse event?

Page x of y

PPD

Programming Note:

- Repeat the same layout for the listing 16.2.7.2
- Population should be used 'Non randomised Subjects'
- The fourth column should be only 'Start Date'
- Delete the footnote related to study day and adjust the numbers accordingly.

CONFIDENTIAL

STATISTICAL ANALYSIS PLAN FOR PROTOCOL 207619

A Clinical Study to Assess the Mildness of a Cosmetic cleanser in Healthy Subjects Using the Forearm-Controlled Application Technique (FCAT)

BIOSTATISTICS DEPARTMENT

GLAXOSMITHKLINE CONSUMER HEALTHCARE

Document type: Statistical Analysis Plan

Authors: PPD (Principal Statistician)

Document status: Amendment Final 1.0

Release date: 26-Jul-2017

Timing of SAP Amendment: ☒ Before unblinding ☐ After un-blinding

Guidance: SAP specific text and bolded (added) or strikethrough (removed) text is an amendment.

Section 2.4.1.2 Statistical hypothesis, model and method of analysis

Reason for amendment: Statistical model of the primary analysis has been updated from Analysis of covariance (ANCOVA) model to Analysis of variance (ANOVA) removing the baseline covariate, since all subjects were only eligible for the trial with dryness and redness scores of zero.

Original text:

The primary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of dryness.

The null and alternative hypotheses for the testing of non-inferiority will be following –

Null Hypothesis (H_{01}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water ≥ 0.25
Alternative Hypothesis (H_{a1}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water < 0.25

To test the above mentioned hypotheses an Analysis of Covariance (ANCOVA) model or appropriate statistical method will be used. The change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; baseline examiner rating of dryness as a covariate and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), and upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of dryness (Table 14.2.2.1).

Amended text:

The primary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of dryness.

The null and alternative hypotheses for the testing of non-inferiority will be following –

Null Hypothesis (H_{01}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water ≥ 0.25

Alternative Hypothesis (H_{a1}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water < 0.25

To test the above mentioned hypotheses an ~~Analysis of Covariance (ANCOVA)~~ Analysis of Variance (ANOVA) model or appropriate statistical method will be used. The change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; ~~baseline examiner rating of dryness as a covariate~~ and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), and upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of dryness (Table 14.2.2.1). If the assumptions underlying the model are strained due to the ordinal nature of the response distribution (or other issues) suitable alternatives may be sought to support the ANCOVA findings.

Section 2.4.1.3 Supportive analysis

Reason for amendment:

The Wilcoxon sign rank test has been removed. The adjusted difference calculation will have the same effect of making all the ranks in one category for one treatment less than the other treatment for that response category, i.e. it would have identical effects on the Wilcoxon test statistic.

Original text:

To support parametric analysis approach described in the Section 2.4.1.2, a one-sided 0.05 level Wilcoxon sign rank test or other statistical method based will be performed on the adjusted differences. The adjusted differences for the test product will be calculated in the following way -

Adjusted differences of the test product at 3 hour on day 5 = Within subject change from baseline of 3 hour on day 5 – 0.25.

The result will be produced in the Table (Table 14.2.2.1).

Frequency and percentage of subjects with clinically relevant irritation responses will be provided by treatment. Clinically relevant irritation response for an individual subject is defined as, increase score in dryness from baseline of 1 or more with treatment (Table 14.2.2.2).

Amended text:

~~To support parametric analysis approach described in the Section 2.4.1.2, a one-sided 0.05 level Wilcoxon sign rank test or other statistical method based will be performed on the adjusted differences. The adjusted differences for the test product will be calculated in the following way -~~

~~Adjusted differences of the test product at 3 hour on day 5 = Within subject change from baseline of 3 hour on day 5 – 0.25.~~

~~The result will be produced in the Table (Table 14.2.2.1).~~

Frequency and percentage of subjects with clinically relevant irritation responses will be provided by treatment. Clinically relevant irritation response for an individual subject is defined as, increase score in dryness from baseline of 1 or more with treatment (Table 14.2.1.1 and 14.2.1.2). Also summary statistics will be provided for dryness and redness by visit and treatment (Table 14.2.1.3 and 14.2.1.4).

Section 2.5.1 Dermal Response (Secondary)

Reason for amendment: Statistical model of the primary analysis has been updated from Analysis of covariance (ANCOVA) model to Analysis of variance (ANOVA) removing the baseline covariate, since all subjects were only eligible for the trial with dryness and redness scores of zero.

Original text:

The first secondary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of redness.

To test the above mentioned hypotheses an Analysis of Covariance (ANCOVA) (ANCOVA) model will be used. The change from baseline in examiner rating of redness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; baseline examiner rating of redness as a covariate and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of redness (Table 14.2.2.2).

Second, secondary endpoint examiner assessment of dryness (Table 14.2.2.3) and redness (Table 14.2.2.4) at day 2,3,4 and 5, 3 hour post last wash procedure will be analyzed using ~~ANCOVA~~ ANOVA model for both dryness and redness separately. The model will be same as mentioned for the primary and first secondary efficacy endpoint analyses. Adjusted means, 95% CI and p-values will be provided for the following treatment comparison.

- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

Change from baseline in TEWL (Table 14.2.3.1) measurement and Corneometer measurement (Table 14.2.3.2) will be analysed using the ANCOVA model with treatment and site as fixed effect; baseline corresponding measurements as a covariate and subject as a random effect. The

adjusted means, treatment differences, 95% CI and p-values will be presented for the following comparisons.

- Test product versus water
- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

If data are sufficiently non-normal in distribution, the non-parametric Wilcoxon signed ranks test or appropriate alternative will be used with median differences and 95% confidence interval presented based on the Hodges-Lehmann method.

No adjustment will be made for multiplicity testing for the secondary endpoints; however consistency of effects will be taken into consideration in interpreting the results.

Amended text:

The first secondary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of redness.

To test the above mentioned hypotheses an Analysis of ~~Covariance~~ Variance (ANOVA) ~~(ANCOVA)~~ model will be used. The change from baseline in examiner rating of redness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; ~~baseline examiner rating of redness as a covariate~~ and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of redness (Table 14.2.2.2).

Second, secondary endpoint examiner assessment of dryness (Table 14.2.2.3) and redness (Table 14.2.2.4) at day 2,3,4 and 5, 3 hour post last wash procedure will be analyzed using ~~ANCOVA~~ ANOVA model for both dryness and redness separately. The model will be same as mentioned for the primary and first secondary efficacy endpoint analyses. Adjusted means, 95% CI and p-values will be provided for the following treatment comparison.

- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

Change from baseline in TEWL (Table 14.2.3.1) measurement and Corneometer measurement (Table 14.2.3.2) will be analysed using the ANCOVA model with treatment and site as fixed effect; baseline corresponding measurements as a covariate and subject as a random effect. The adjusted means, treatment differences, 95% CI and p-values will be presented for the following comparisons.

- Test product versus water
- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

If data are sufficiently non-normal in distribution, the non-parametric Wilcoxon signed ranks test or appropriate alternative will be used with median differences and 95% confidence interval presented based on the Hodges-Lehmann method.

No adjustment will be made for multiplicity testing for the secondary endpoints; however consistency of effects will be taken into consideration in interpreting the results.

Section 3 Changes to the Protocol Defined Statistical Analysis Plan

Reason for amendment: Changes have been made in the Statistical Analysis Plan which were not planned for in the Protocol.

Original text:

There were no changes or deviations to the originally planned statistical analysis specified in the protocol version 1.0 [(Dated: 06/MAR/2017)].

Amended text:

~~There were no changes or deviations to the originally planned statistical analysis specified in the~~ The use of the Wilcoxon type test in support of the primary was removed due to limitations of that approach. Any potential supportive analysis will be considered if the ANOVA assumptions are not met. protocol version 1.0 [(Dated: 06/MAR/2017)].

Section 5.1 Templates for Tables

Reason for amendment: Other minor changes to the templates have been included which includes splitting of dryness and redness displays.

The purpose of this Statistical Analysis Plan (SAP) is to describe the planned analyses and output to be included in the Clinical Study Report (CSP) for Protocol 207619. This SAP will be finalised prior to database freeze and treatment code un-blinding.

1 Study details

A cosmetic product that is freely available to the consumer must be safe when applied under normal or reasonably foreseeable conditions of use (ANVISA, 2012). Thus, the raw materials used in the product formulation must be of proven safety and with established use in the cosmetic industry. As a general requirement, the safety of the final formulation must also be confirmed before it is marketed.

Cleansers are designed to remove dirt, sweat, sebum and oils from the skin. This is achieved through the use of surfactants that aid in the uplifting of dirt and solubilisation of oils. However the interaction of cleansers and the stratum corneum can be detrimental to the skin, causing tightness and dryness as well as barrier damage, erythema, irritation and itch. (Wihelm, 1994)

Mildness is a factor that contributes to consumer acceptance of a personal cleansing product, especially in dry skin. The relative mildness (or irritation potential) of personal cleansing products is ideally judged under conditions of actual consumer use. However, these may not present the best conditions for discriminating product mildness differences, since normal use conditions typically do not induce differentiable product skin effects (Ertel, 1995).

The Forearm Controlled Application Technique (FCAT) is an exposure method that offers an efficient means to estimate the relative mildness (irritation) of personal cleansers through repeated overuse. It provides greater precision and sensitivity and minimises confounding effects due to biological diversity (Ertel, 1995). The FCAT is able to discriminate the relative mildness of personal cleansers in a wide variety of test conditions. Mildness can be defined by the parameters of skin redness and dryness (Lukacocvic, 1987).

Transepidermal water loss (TEWL) is the rate at which water permeates the stratum corneum and evaporates from the skin surface and qualifying TEWL can be used to support the examiner assessments.

Corneometry is a standard measurement of skin moisturisation which uses an electrical capacitance method to model the water content of the skin (Eisner et al, 1994). This assessment is included to provide an instrumental measurement of skin moisturisation to support the examiner assessments.

The objective of this clinical study is to assess the relative mildness of a cosmetic facial cleanser in comparison to water through repeated application to the volar forearm using the FCAT wash procedure. A positive control has been included in the design to help validate the trial results. Specifically, Imperial Leather Original bar soap has been selected as the positive control as it has shown to have induced greater skin dryness than water in a previous study CCI [redacted]

General safety and tolerability will be assessed based on the frequency and severity of Adverse Events (AEs).

1.1 Study design

Overall Design

This is a test site randomised, examiner blinded, positive and negative-controlled, single-center; Forearm Controlled Application Technique clinical study in healthy subjects to assess the mildness potential of a cosmetic facial cleansing product.

Subjects will give their written consent prior to any study procedures taking place. Visual assessments of dryness and redness will be conducted (by a trained blinded examiner). Each subject should have a score of zero for dryness and zero for redness on each volar forearm to be eligible for inclusion in the study. Eligible subjects will undergo a 5 to 7-day washout period, during which only the provided standard soap (Simple Soap®) will be used. This standard soap will also be used throughout the study in place of current cleansing product(s), but subjects should avoid using anything other than water on the volar forearms throughout the study, including the washout period.

A qualified dermatologist will assess subjects at Screening (Visit 1) for eligibility and again at Visit 2 for continued eligibility, to ensure the subjects are free of any clinically-relevant dermatological conditions. At the Baseline Visit (Visit 2, Day 1) examiner assessment of the dryness and redness of each proposed test site on each volar forearm will be performed prior to any wash application using the scoring system detailed in Appendix 2 of the Protocol. Each subject should have a score of zero for dryness and zero for redness to be considered eligible to continue in the study.

Prior to any wash procedure, subjects will be acclimatized to the environment of controlled temperature and humidity room and instrumental measurements of skin barrier function (measured with a Tewameter) and skin moisturisation (measured with a Corneometer) will be performed at Baseline (Visit 2 (AM)) in addition to the Baseline examiner ratings of dryness and redness at each forearm test site.

This will be followed by 5 days of repeated FCAT wash procedure where subjects will be instructed to attend the site for two wash visits per day for the first 4 days (Visits 2-5, AM and PM) and one wash visit on Day 5 (Visit 6, AM only). The AM and PM wash procedures on any given day will be separated by a minimum of 3 hours.

Visual assessment of dryness and redness will be conducted (by a trained, blinded examiner) prior to each PM wash procedure for Days 1-4 (Visits 2-5). The final examiner assessment of dryness and redness will be conducted on Day 5 (Visit 6), a minimum of 3 hours following the final Day 5 AM wash procedure.

A trained technician will perform each wash procedure to ensure consistency and eliminate the risk of cross-contamination of the study test sites. The technician will wash each of the

4 designated sites with one of the randomly allocated products; test product, positive control (Imperial Leather bar soap), negative control (sterile water only) or unwashed.

On Day 5, 3 hours after the last wash procedure, examiner ratings of dryness and redness and final instrumental measurements of skin barrier function and skin moisturisation will be performed. Additionally, a final clinical assessment by a qualified dermatologist will be performed to ensure that it is medically appropriate to discharge each subject from the study.

1.2 Study objectives

Objectives	Endpoints
Primary Objective	**Primary Endpoint**
• To compare change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5, for the test product versus water in the skin of healthy subjects.	• Examiner assessment of dryness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hours post last wash procedure.
Secondary Objectives	**Secondary Endpoint**
• To compare change from baseline in the examiner rating of redness to 3 hours following last wash procedure on day 5 for the test product versus water in the skin of healthy subjects.	• Examiner assessment of redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hours post last wash procedure.
• To compare change from baseline in the examiner rating of dryness and redness each day for test product and the positive control versus water in the skin of healthy subjects.	• Examiner assessment of dryness and redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 2, 3, 4 and 5 and 5, 3 hours post last wash procedure.
• To assess change from baseline in skin barrier function following the FCAT wash procedure for test product and positive control versus water.	• TEWL measurements on day 5, 3 hours post last wash procedure.
• To assess change from baseline skin moisturisation following the FCAT wash procedure for test product and positive control versus water.	• Corneometer measurements on day 5, 3 hours post last wash procedure.
• To evaluate the general safety of the test product.	• Assessment of frequency and severity of Adverse Events.

1.3 Treatments

	Test Product	Positive Control	Reference Product
Product Name	Micellar cleanser	Imperial Leather Original Bar Soap	Baxter Sterile Water

Product Formulation Code (MFC)	CCI [redacted]	N/A Market Place product	N/A Market Place product
Product Format	200 ml clear PET Bottle	100g Bar	1000ml Bottle
Application Quantity	0.09ml on to a moistened towel (with sterile water)	Moistened Towel(with sterile water) rubbed onto bar soap for 6 seconds to generate a lather	0.09ml on to a moistened towel (with sterile water)
Route of Administration	Topical dermal application		
Application Instruction	Applied on-site by technician		

Each subject will have four test sites; two test sites (3x3 cm) marked out on each volar forearm (Figure 1, Section 3.4). The distance between different application sites must be at least 1 cm to prevent either the test product, positive control (Imperial Leather bar soap) or negative control (sterile water) wash from spreading over and influencing neighbouring test sites. The site on each forearm for each wash application (or unwashed) will be determined according to the provided randomisation schedule. The randomisation schedule will identify the test sites for the test product, positive control and water (negative control) wash applications, and the unwashed site.

A trained technician will wash each marked site with the appropriate product.

All wash procedures will be performed at the study site by a technician to ensure consistent wash application and correct product usage. During the study, subjects will have the wash procedure completed four times daily at each allocated site, two wash applications in the morning (AM) and two wash applications in the afternoon (PM) on Days 1, 2, 3 and 4. On Day 5, subjects will have the wash procedure completed twice in the morning (AM) only. Each AM and PM wash procedure will be separated by a minimum of 3 hours.

1.4 Time points and visit windows

All data will be accepted for the analysis. Deviations from the scheduled assessment times are expected to be small and few. The following are the acceptable time windows.

Day#/Visit#	Activity	Time window
Day 1 (AM)/Visit 2	Wash procedures	At least 3 hours apart
Day 1 (PM)/Visit 2	Wash procedures	At least 3 hours apart
Day 2 (AM)/Visit 3	Wash procedures	At least 3 hours apart

Day#/Visit#	Activity	Time window
Day 2 (PM)/Visit 3	Wash procedures	At least 3 hours apart
Day 3 (AM)/Visit 4	Wash procedures	At least 3 hours apart
Day 3 (PM)/Visit 4	Wash procedures	At least 3 hours apart
Day 4 (AM)/Visit 5	Wash procedures	At least 3 hours apart
Day 4 (PM)/Visit 5	Wash procedures	At least 3 hours apart
Day 5 (AM)/Visit 6	Wash procedures	At least 3 hours apart

During screening subjects will sign an informed consent document and then a dermatological assessment will be conducted to ensure subjects have no dermatological conditions on their forearms that might impact subject safety or study results and to ensure subjects are classified as Fitzpatrick Phototype I to IV. A trained, blinded examiner will assess the dryness and redness of each volar forearm and only subjects with a score of zero for dryness and zero for redness will be enrolled into the study. Each subject's medical and medication history will be reviewed, in addition to the inclusion/exclusion criteria. Subsequently, site staff will review lifestyle guidelines and directions with eligible subjects.

A standard soap to be used in place of subjects' normal cleansers will be provided for use during the washout period and study treatment phase. Subjects will be instructed to avoid using any products, including the supplied cleanser, on their forearms throughout the washout and study treatment phase.

Consented eligible subjects will undergo a one-week washout period (5-7 days) prior to first wash procedure, during which the use of any skin cleansing products, including the standard soap, and any other topical products on the forearms will be prohibited.

On Day 1 (Visit 2, AM), eligible subjects will return to the study site and another dermatological assessment will be conducted, as will a review of any concomitant medications used since screening to ensure subject continued eligibility. Compliance concerning correct use of the standard soap will also be checked. Each subject will have two test sites (3x3 cm) marked out on each volar forearm. The distance between the two sites should be at least 1 cm to prevent either the test product, positive control product (Imperial Leather Original bar soap) or negative control (sterile water) applications from spreading over and influencing neighbouring test sites. A trained, blinded examiner assessment (Baseline; Day 1 - AM) of dryness and redness will be conducted at each marked site to ensure the dryness score is zero and the redness score is zero for each site for continued eligibility. The site on each forearm for each wash procedure (or unwashed) will be determined according to the provided randomisation schedule.

The randomisation schedule will determine the allocation of the 4 test sites for the test product, positive control and negative control wash applications and the site to be left unwashed. A trained technician will perform product application to ensure consistency and eliminate the risk of cross contamination of the study test sites.

Subjects will be acclimatised to the environment of a temperature (20-22ºC) and humidity (40-60% RH) controlled room for a period of at least 30 minutes before instrumental measurements of TEWL and corneometry at Baseline (Day 1 - AM) are taken, prior to any wash procedures and additionally at the final visit (Day 5 - PM) following the final wash procedure (Day 5 – AM). (EEMCO, 1997).

The treatment period will consist of 5 days of FCAT wash procedure where subjects will be instructed to come to the site for two wash procedures per day for the first 4 days (AM and PM) and one wash procedure on Day 5 (AM only). Each wash procedure will consist of two washings of each test site AM and PM. The AM and PM wash procedures on any given day will be separated by a minimum of 3 hours.

Further visual assessments of dryness and redness will be performed prior to each PM wash procedure on Days 1 to 4 and again at Day 5 (PM) following the final Day 5AM wash procedure (Appendix 2). The visual assessments of dryness and redness will be performed prior to subject acclimatisation, TEWL and Corneometry measurements on Day 1 and Day 5.

The FCAT generally produces only mild to moderate skin irritation, however, if any site reaches a dryness or redness score of ≥ 5.0 at any time during the study, treatment of all sites on that subject will be immediately discontinued, all test sites will be scored and these will be recorded as an Adverse Event and the subject will be withdrawn from the study.

2 Data analysis

Data analysis will be performed by Biostatistics Division in inVentiv International Pharma Services. Prior to database closure a Blind Data Review Meeting (BDRM) will be conducted in which various aspects of the trial will be discussed and agreed. The statistical analysis software used will be SAS Studio version 9.4 in a WINDOWS environment.

Except as described below, all listings will be produced for all randomised subjects.

2.1 Populations for analysis

Tables described in this section will be produced for all randomised subjects unless it is not mentioned otherwise in the specified sub-section.

2.1.1 Subject disposition

Subject disposition will be summarized as the number and percentage of subjects who complete the study, with the number who discontinue broken down by reason for discontinuation (Table 14.1.1) using all screened subjects.

2.1.2 Protocol Deviations

Important protocol violations (including violations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarised and listed.

Protocol violations will be tracked by the study team throughout the conduct of the study. Data will be reviewed prior to un-blinding and closure of the database to ensure all important violations are captured and categorised.

Major violations of the protocol procedures identified as liable to influence the efficacy outcome are as follows:
All-important violations will be defined in the "Review Listing Requirement (RLR)" document.

A list of protocol deviation will be listed in Listing 16.2.2.

2.1.3 Analysis populations

Four analysis populations are defined.

Population	Definition / Criteria	Analyses Evaluated
All Screened Subjects	• All subjects those who are screened	• Disposition
Randomised	• Randomised population will include all subjects who may or may not receive the any wash procedure.	• Protocol violation
Safety	• Safety population includes all subjects who received at least one wash procedure.	• Safety Analyses
Intent-to-Treat (ITT)	• The 'Intent to treat' (ITT) population includes all subjects who are randomised, receive at least one wash procedure, including sterile water, and have at least one post-baseline assessment available.	• Efficacy analysis (Dryness, Redness, etc.)

2.1.4 Subgroups/Stratifications

Not Applicable.

2.1.5 Centres pools

Not Applicable.

2.2 Patient demographics/other Baseline characteristics

Demographic and baseline characteristics summaries will be produced for the safety and ITT.

2.2.1 Demographic and Baseline characteristics

Categorical demographic variables include gender, race and Fitzpatrick scale. These variables will be summarized by the number and percentage of subjects with each relevant characteristic

(Table 14.1.4.1 and Table 14.1.4.2). Age will be summarized by the number of non-missing observations, mean, and standard deviation, median, minimum and maximum values.

The Fitzpatrick scale is a numerical classification that is widely used by dermatologists to classify a person's skin type by their response to the sun exposure (Fitzpatrick, 1988).

CCI

[redacted]

2.2.2 General medical history

Medical history data will not be presented in the study report. A data listing will be produced for evaluation of protocol violations only at the blinded data review stage.

2.2.3 Characteristics of Disease

Not Applicable.

2.3 Treatments (study treatment, other concomitant therapies, compliance)

Compliance data will be summarized for the ITT population. Exposure and other medications will be summarized on the safety population.

2.3.1 Study Product/treatment Compliance and Exposure

Not Applicable.

2.3.2 Any protocol deviation associated with bar soap wash adherence will be listed at the blinded data review stage. Concomitant medication

Concomitant medication/non-drug treatments data will not be presented in the study report. A data listing will be produced for evaluation of protocol violations only at the blinded data review stage.

2.4 Analysis of dermal response

2.4.1 Primary dermal response endpoint

2.4.1.1 Primary dermal response endpoint definition

The primary analysis will be performed to compare change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 for the test product versus water in the skin of healthy subjects. Analyses will be performed on the ITT population.

The forearm grading scale is described below –

Forearm Grading Scale:
Any skin response at any of the wash procedure areas on the volar forearm will be clinically assessed using the criteria recommended by Lukacovic, 1987. The following grades will be used to express the response observed at the time of examination:

CCI

CCI

CCI

Individual subjects observations for dryness and redness scores, TEWL and corneometer will be listed by visit and treatment (Listing 16.2.6).

Frequency table of dryness and redness score measurements at each time point will be presented based of ITT population (Table 14.2.1.1, 14.2.1.2). Separate summary statistics table (displaying number of non-missing observations, mean, standard deviation, median, minimum, and maximum) of the raw and change from baseline in dryness and redness scores, TEWL and corneometer measurements at each visit will be summarized for test product, water, positive control and unwashed site (Table 14.2.1.3, 14.2.1.4, 14.2.1.5 and 14.2.1.6) using ITT population.

2.4.1.2 Statistical hypothesis, model, and method of analysis

The primary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of dryness.

The null and alternative hypotheses for the testing of non-inferiority will be following –

Null Hypothesis (H_{01}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water ≥ 0.25
Alternative Hypothesis (H_{a1}):	The population mean of dryness (in the examiner rating scale) for test product minus the population mean of dryness (in the examiner rating scale) for water < 0.25

To test the above mentioned hypotheses an ~~Analysis of Covariance (ANCOVA)~~ Analysis of Variance (ANOVA) model or appropriate statistical method will be used. The change from baseline in examiner rating of dryness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; ~~baseline examiner rating of dryness as a covariate~~ and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), and upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of dryness (Table 14.2.2.1). If the assumptions underlying the model are strained due to the ordinal nature of the response distribution (or other issues) suitable alternatives may be sought to support the ANCOVA findings.

2.4.1.3 Supportive analyses

~~To support parametric analysis approach described in the Section 2.4.1.2, a one-sided 0.05 level Wilcoxon sign rank test or other statistical method based will be performed on the adjusted differences. The adjusted differences for the test product will be calculated in the following way~~

~~Adjusted differences of the test product at 3 hour on day 5 = Within subject change from baseline of 3 hour on day 5 – 0.25.~~

~~The result will be produced in the Table (Table 14.2.2.1).~~

Frequency and percentage of subjects with clinically relevant irritation responses will be provided by treatment. Clinically relevant irritation response for an individual subject is defined as, increase score in dryness from baseline of 1 or more with treatment (Table 14.2.1.1 and 14.2.1.2).

2.4.2 Secondary dermal response variables

The secondary efficacy variables are defined in the following sub-sections.

2.4.2.1 Secondary dermal response variable 1

The first secondary efficacy variable is examiner assessment of redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 5, 3 hour post last wash procedure.

2.4.2.2 Secondary dermal response variable 2

- The secondary efficacy variables are the examiner assessment of dryness and redness (Lukacovic, 1987) as per Appendix 2 of the protocol at day 2, 3 4 and 5, 3 hour post last wash procedure.
- Also a secondary efficacy variable is TEWL measurements on day 5, 3 hours post last wash procedure.
- Also a secondary efficacy variable is Corneometer measurements on day 5, 3 hours post last wash procedure.

2.4.3 Handling of missing values/censoring/discontinuations

Missing data will not be replaced or imputed. Dropouts will be included in analyses up to the point of discontinuation. Sensitivity analyses may be conducted if substantial number of subjects have discontinued from the study.

2.5 Analysis of secondary objectives

2.5.1 Dermal Response (secondary)

The first secondary endpoint will be tested for the non-inferiority of the test product to water in change from baseline on day 5 in examiner rating of redness.

To test the above mentioned hypotheses an Analysis of ~~Covariance~~ Variance (ANOVA) ~~(ANCOVA)~~ model will be used. The change from baseline in examiner rating of redness to 3 hours following last wash procedure on day 5 will be considered as the response variable. The model will include treatment and site as fixed effects; ~~baseline examiner rating of redness as a covariate~~ and subject as a random effect.

Least Square Means (LSMeans) for each treatment will be presented. The treatment difference estimates (test product to water), upper one-sided 90% confidence limit will be presented. If the upper one-sided 90% confidence limit is less than 0.25 it will be concluded that the test product is not inferior to the water in the examiner rating of redness (Table 14.2.2.3).

Second, secondary endpoint examiner assessment of dryness (Table 14.2.2.4) and redness (Table 14.2.2.5)at day 2,3,4,and 5, 3 hour post last wash procedure will be analyzed using ~~ANCOVA~~ ANOVA model for both dryness and redness separately. The model will be same as mentioned for the primary and first secondary efficacy endpoint analyses. Adjusted means, 95% CI and p-values will be provided for the following treatment comparison.

- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

Change from baseline in TEWL (Table 14.2.3.1) measurement and Corneometer measurement (Table 14.2.3.2) will be analysed using the ANCOVA model with treatment and site as fixed effect; baseline corresponding measurements as a covariate and subject as a random effect. The adjusted means, treatment differences, 95% CI and p-values will be presented for the following comparisons.

- Test product versus water
- Positive Control versus water
- Positive Control versus test product
- Water versus un-washed area

If data are sufficiently non-normal in distribution, the non-parametric Wilcoxon signed ranks test or appropriate alternative will be used with median differences and 95% confidence interval presented based on the Hodges-Lehmann method.

No adjustment will be made for multiplicity testing for the secondary endpoints; however consistency of effects will be taken into consideration in interpreting the results.

2.5.2 Safety

2.5.2.1 Adverse events and Serious Adverse Events

All adverse events (AEs) will be listed in the (Listing 16.2.7.1) for the randomised subjects and all adverse events for the non-randomised subjects will be listed in Listing 16.2.7.2.

Treatment emergent adverse events (TEAEs) will be summarized by the number and percentage of subjects having any adverse event, an adverse event in each System Organ Class, and each individual adverse event (Table 14.3.1.1). All TEAEs will also be tabulated by severity and treatment (Table 14.3.1.2). Treatment-emergent AEs suspected of a relationship to study medication and those causing study discontinuation will be presented in a similar manner (Table 14.3.1.3). For treatment-related AEs, these will also be presented by severity and treatment, if applicable (Table 14.3.1.4).

Non-fatal serious adverse events and adverse events causing study treatment discontinuation will be listed (Listing 14.3.2.1).

2.6 Analysis of other variables

Not applicable.

2.7 Interim analysis

No interim analysis is planned.

2.8 Sample size calculation

The primary hypothesis is that the test product is not inferior to water in the change from baseline in the examiner rating of dryness on Day 5. Assuming that the standard deviation of the within subject treatment difference in change from baseline in examiner rating of dryness is 0.25 (based on previous study GSKCH: CCI [redacted]

A sufficient number of subjects will be screened in order to randomise 45 subjects and to ensure that at least 40 subjects complete the study.

3 Changes to the Protocol Defined Statistical Analysis Plan

~~There were no changes or deviations to the originally planned statistical analysis specified in the protocol version 1.0~~ The use of the Wilcoxon type test in support of the primary was removed due to limitations of that approach. Any potential supportive analysis will be considered if the ANOVA assumptions are strained. protocol version 1.0 [(Dated: 06/MAR/2017)].

4 Appendix 1:

4.1 List of Tables, Listings and Figures

4.2 Tables

Table Number	Table Title (Population)	Template
14.1.1	Subject Disposition (All Screened Subjects)	Appendix 2
14.1.4.1	Subject Demographics and Baseline Characteristics (Safety Population)	Appendix 2
14.1.4.2	Subject Demographics and Baseline Characteristics (ITT Population)	14.1.4.1
14.2.1.1	Frequency of Examiner Assessment Rating on Dryness by Visit and Treatment (ITT Population)	Appendix 2
14.2.1.2	Frequency of Examiner Assessment Rating on Redness by Visit and Treatment (ITT Population)	14.2.1.1
14.2.1.3	Summary of Examiner Assessment Rating on Dryness by Visit and Treatment (ITT Population)	Appendix 2
14.2.1.4	Summary of Examiner Assessment Rating on Redness by Visit and Treatment (ITT Population)	14.2.1.3
~~14.2.1.1~~	~~Summary of Examiner assessment rating on dryness and redness by Visit and Treatment (ITT Population)~~	~~Appendix 2~~
14.2.1.5	Summary of TEWL ((g/m^2/hr))~~and Corneometer~~ measurements by Visit and Treatment (ITT Population)	14.2.1.3
14.2.1.6	Summary of Corneometer (i.u.) Measurements by Visit and Treatment (ITT Population)	14.2.1.3
14.2.2.1	Statistical Analysis of Examiner Assessment of Dryness on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	Appendix 2
~~14.2.2.2~~	~~Summary of clinically relevant examiner assessment of dryness and redness on day 2,3,4,and 5, 3 hour post last wash procedure by treatment (ITT Population)~~	~~Appendix 2~~
14.2.2.2	Statistical Analysis of Examiner Assessment of Redness on Day 5, 3 hour Post Last Wash Procedure (ITT Population)	14.2.2.1

Table Number	Table Title (Population)	Template
14.2.2.3	Statistical Analysis of Examiner Assessment of Dryness on Day 2,3,4 and 5, 3 Hhour Post Last Wash Procedure (ITT Population)	Appendix 2
14.2.2.4	Statistical Analysis of Examiner Assessment of Redness on Day 2,3,4 and 5, 3 Hour Post Last Wash Procedure (ITT Population)	14.2.2.3
14.2.3.1	Statistical Analysis of TEWL Measurement on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	14.2.2.3
14.2.3.2	Statistical Analysis of Corneometer Measurement on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	14.2.2.3
14.3.1.1	Treatment Emergent Adverse Event (Safety Population)	Appendix 2
14.3.1.2	Treatment Emergent Adverse Event by Severity and Treatment (Safety Population)	Appendix 2
14.3.1.3	Treatment Emergent Treatment Related Adverse Event (Safety Population)	14.3.1.1
14.3.1.4	Treatment Emergent Treatment Related Adverse Event by Severity and Treatment (Safety Population)	14.3.1.2

4.3 Listings

Listing Number	Listing Title (Population)	Template
14.3.2.1	Listing of Serious Adverse Events Leading to Discontinuation (Randomised Population)	16.2.7.1
14.3.2.2	Listing of Deaths (Randomised Population)	16.2.7.1
16.1.7.	Randomization Information (Randomised Population)	Appendix 2
16.1.9.1	Statistical Analysis of Examiner Assessment of Dryness on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis
16.1.9.2	Statistical Analysis of Examiner Assessment of Redness on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis
16.1.9.3	Statistical Analysis of Examiner Assessment of Dryness on Day 2,3,4 And 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis

Listing Number	Listing Title (Population)	Template
16.1.9.4	Statistical Analysis of Examiner Assessment of Redness on Day 2,3,4 and 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis
16.1.9.5	Statistical Analysis of TEWL Measurement on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis
16.1.9.6	Statistical Analysis of Corneometer Measurement on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)	SAS output for model analysis
16.2.2	Individual Subjects Protocol Deviations (Randomised Population)	Appendix 2
16.2.6	Subjects with Skin Dryness, Redness Rating and Measurement on TEWL and Corneometer (ITT Population)	Appendix 2
16.2.7.1	Listing of All Adverse Events (Randomised Population)	Appendix 2
16.2.7.2	Listing of All Adverse Events (Non-randomised Subjects)	16.2.7.1

Note: If there are no data to display generate a null listing.

4.4 Top line Outputs:

Table/Listing Figure Number	Table/Listing/Figure Title (Population)
14.1.1.1	Subject Disposition (All Screened Subjects)
14.1.4.2	Subject Demographics and baseline Characteristics (ITT Population)
14.2.2.1	Statistical Analysis of Examiner Assessment of Dryness on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)
14.2.2.2	Statistical Analysis of Examiner Assessment of Redness on Day 5, 3 Hour Post Last Wash Procedure (ITT Population)
14.3.1.1	Treatment Emergent Adverse Event (Safety Population)
16.2.7.1	Listing of All Adverse Events (Randomised Population)

5 Appendix 2:

5.1 Templates for the Tables, listings and figures

This is a guideline which will give the guidance of treatment labels that will be used for the table header and in the figures, listings and in the footnotes.

The treatment labels for the column heading will be as follow:

- Micellar cleanser
- Imperial Leather Original Bar Soap
- Baxter Sterile Water
- Unwashed

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.1.1
Subject Disposition
All Screened Subjects

All Screened Subjects (N=XX)

	Overall(N=XX) n (%)
TOTAL NUMBER OF SUBJECTS SCREENED	xx (xx.x)
SUBJECTS NOT RANDOMISED	xx (xx.x)
DID NOT MEET STUDY CRITERIA	xx (xx.x)
ADVERSE EVENTS	xx (xx.x)
ETC.	xx (xx.x)
SUBJECTS RANDOMISED	xx (xx.x)
COMPLETED	xx (xx.x)
DID NOT COMPLETE	xx (xx.x)
ADVERSE EVENT	xx (xx.x)
LOST TO FOLLOW UP	xx (xx.x)
PROTOCOL DEVIATION	xx (xx.x)
WITHDRAWAL OF CONSENT	xx (xx.x)
OTHER	xx (xx.x)
RANDOMISED POPULATION	xx (xx.x)
SAFETY POPULATION	xx (xx.x)
INTENT TO TREAT POPULATION	xx (xx.x)

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.1.4.1
Subject Demographics and Baseline Characteristics
Safety Population

Safety Population (N=XX)

Demographic variables	Overall (N=XX)
SEX n (%)	
MALE	xx (xx.x)
FEMALE	xx (xx.x)
RACE n (%)	
ASIAN	
BLACK OR AFRICAN AMERICAN	
AMERICAN INDIAN OR ALASKA NATIVE	xx (xx.x)
NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	xx (xx.x)
WHITE	xx (xx.x)
MULTIPLE	xx (xx.x)
AGE (YEARS)	
N	XX
MEAN	XX.X
SD	XX.XX
MEDIAN	XX.X
MINIMUM	XX
MAXIMUM	XX
FITZPATRICK SCALE FOR SKIN TYPE	
I	xx (xx.x)
II	xx (xx.x)

III	xx (xx.x)
IV	xx (xx.x)
V	xx (xx.x)
VI	xx (xx.x)

Fitzpatrick Scale For Skin Type: I= Always burns easily; never tans (pale white skin);II= Always burns easily; tans minimally (white skin);III= Burns moderately; tans gradually (light brown skin); IV= Burns minimally, always tans well (moderate brown skin);V= Rarely burns, tans profusely (dark brown skin); VI= Never burns (deeply pigmented dark brown to black skin)

PPD

Programming Note:

- Repeat Table 14.1.4.2 with ITT Population.
- Also, for the corneometer assessment

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.1.1
Frequency of Examiner Assessment Rating on Dryness by Visit and Treatment
Intent to Treat Population

Study Population: ITT Population (N=XX)

	Micellar Cleanser (N=XX)	Imperial Leather Original Bar Soap (N = xx)	Baxter Sterile Water (N = xx)	Unwashed (N=xx)
Day/Visit	n (%)	n (%)	n (%)	n (%)
Baseline				
Score=0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Day 1 (AM)/Visit 2				
Total	xx	xx	xx	xx
Missing	xx	xx	xx	xx
Score=0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=0.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=1	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=1.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=2	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=2.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=3	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=3.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=4	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=4.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=5.5	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Score=6	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Clinically Relevant Irritation Response [1]	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Day 1 (PM)/Visit 2				
Day 2 (PM)/Visit 3				
Day 5 (PM)/Visit 6				
Day 5 (PM)/Visit 6				

[1] Clinically relevant irritation response is defined as an increase score in dryness from baseline of 1 or more.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.1.3
Summary of Examiner Assessment Rating on Dryness by Visit and Treatment
Intent to Treat Population

Study Population: ITT Population (N=XX)

	Micellar Cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
Day (AM/PM)/Visit	Raw	Change from bl.	Raw	Change from bl.	Raw	Change from bl.	Raw	Change from bl.
Day 1 (AM)/Visit 2								
n	xx	xx	xx	xx	xx	xx	xx	xx
Missing	xx	xx	xx	xx	xx	xx	xx	xx
Mean	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
SE	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
Minimum	xx	xx	xx	xx	xx	xx	xx	xx
Maximum	xx	xx	xx	xx	xx	xx	xx	xx
Day 1 (PM)/Visit 2								
Day 2 (PM)/Visit 3								
……..								
Day 5 (PM)/Visit 6								
Day 1 (AM)/Visit 2								
……..								
Day 5 (PM)/Visit 6								

Table 14.2.1.1
Summary of Examiner Assessment rating on Dryness, Redness by Visit and Treatment
Intent to Treat Population

Intent to Treat Population N=XX)

Day/Visit/Assessment		Micellar cleanser (N=XX)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
		Raw	Change	Raw	Change	Raw	Change	Raw	Change
Baseline	N								
	SCORE=0	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
Day 1 (AM)/Visit 2/Dryness	N	xx		xx		xx		xx	
	MISSING	xx		xx		xx		xx	
	SCORE=0	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=1	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=2	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=3	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=4	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=5	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	SCORE=6	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
	MEAN	xx.xx		xx.xx		xx.xx		xx.xx	
	SD	xx.xxx		xx.xxx		xx.xxx		xx.xxx	
Day 1 (PM)/Visit 2/Dryness	Same as above								
Day 2 (PM)/Visit 3/Dryness	Same as above								
……									
Day 5 (PM)/Visit 6/Dryness	Same as above								
Day 1 (AM)/Visit 2/Redness	Same as above								
……									

Day/Visit/Assessment		Micellar cleanser (N=XX)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
		Raw	Change	Raw	Change	Raw	Change	Raw	Change
Baseline	N								
	SCORE=0	xx (xx.x)		xx (xx.x)		xx (xx.x)		xx (xx.x)	
Day 1 (AM)/Visit 2/Dryness	N	xx		xx		xx		xx	
Day 5 (PM)/Visit 6/Redness	Same as above								

Dryness rating: 0= No dryness; 1.0=Patches of slight powderiness and occasional patches of small scales may be seen; 2.0= Generalised slight powderiness; minimal popular response; 3.0= Generalised moderate powderiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powderiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness; 6.0=Severe redness.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.1.2
Summary of TEWL (g/m²/hr) and Corneometer (i.u.) measurements by Visit and Treatment
Intent to Treat Population

Intent to Treat Population N=XX)

Day (AM/PM)/Visit/Assessment	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	Raw	Change	Raw	Change	Raw	Change	Raw	Change
Day 1 (AM)/Visit 2/TEWL								
n	xx		xx		xx		xx	
Missing	xx		xx		xx		xx	
Mean	xx.x		xx.x		xx.x		xx.x	
SD	xx.xx		xx.xx		xx.xx		xx.xx	
SE	xx.xx		xx.xx		xx.xx		xx.xx	
Median	xx.x		xx.x		xx.x		xx.x	
Minimum	xx		xx		xx		xx	
Maximum	xx		xx		xx		xx	

Page x of y

PPD

Programming Note:

- Repeat for Day 5
- Also, for the corneometer assessment

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.2.1
Statistical Analysis of Examiner Assessment of Dryness on Day 5, 3 Hour Post Last Wash Procedure
Intent to Treat Population

Intent to Treat Population N=XX)

	Micellar cleanser (N = xx)		Baxter Sterile Water (N = xx)	
	Baseline [1]	Change	Baseline [1]	Change
Day 5, 3 hour post last wash				
n	xx	xx	xx	xx
Missing	xx	xx	xx	xx
Mean	xx.x	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx	xx.xx
SE	xx.xx	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x	xx.x
Minimum	xx	xx	xx	xx
Maximum	xx	xx	xx	xx
Adjusted Mean (SE) [2]		xx.x (xx.xx)		xx.x (xx.xx)
90% CI [2]		(xx.x, xx.x)		(xx.x, xx.x)
P-value [2]		0.xxxx		0.xxxx

Comparison Between Treatments	Difference [2,3]	Upper One-sided 90% CI	P-value
Micellar cleanser vs Baxter Sterile Water	xx.x	(xx.x, xx.x)	0.xxxx

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.
[2] From ANOVA model with change from baseline in examiner assessment on dryness as response and treatment, site as fixed effects. If one-sided upper 95% CI is less than 0.25 then non-inferiority is established.
[3] Difference is first named treatment minus second named treatment such that a negative difference favours first named treatment.
[4] P-value from Wilcoxon sign rank test, which is based on the adjusted differences. Adjusted difference is calculated in the following manner;

Table 14.2.2.2
Summary of clinically relevant examiner assessment of dryness and redness on day 2,3,4,and 5, 3 hour post last wash procedure by treatment
Intent to Treat Population

Intent to Treat Population (N=XX)
Treatment:Miscellar Cleanser

Day # (time)/Visit #	Rating	Dryness Rating Baseline							Redness Rating Baseline						
		0	1.0	2.0	3.0	4.0	5.0	6.0	0	1.0	2.0	3.0	4.0	5.0	6.0
Day 1(PM)/Visit 2	0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	1.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	2.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	3.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	4.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	5.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
	6.0	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)	xx (xx.x)
Day 2 (PM)/Visit 3															
…….															
Day 5 (PM)/Visit 6															

Dryness rating: 0= No dryness; 1.0=Patches of slight powderiness and occasional patches of small scales may be seen; 2.0= Generalised slight powderiness; minimal popular response; 3.0= Generalised moderate powderiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powderiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness;6=Severe redness.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.2.2.3
Statistical Analysis of Examiner Assessment of Dryness on Day 2, 3, 4 and 5, 3 Hour Post Last Wash Procedure
Intent to Treat Population

Intent to Treat Population N=XX)

	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	Baseline [1]	Change	Baseline [1]	Change	Baseline [1]	Change	Baseline [1]	Change
Day 2, 3 hour post last wash								
n	xx	xx	xx	xx	xx	xx	xx	xx
Missing	xx	xx	xx	xx	xx	xx	xx	xx
Mean	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
SE	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
Minimum	xx	xx	xx	xx	xx	xx	xx	xx
Maximum	xx	xx	xx	xx	xx	xx	xx	xx
Adjusted Mean (SE) [2]		xx.x (xx.xx)		xx.x (xx.xx)	xx.x (xx.xx)		xx.x (xx.xx)	
95% CI [2]		(xx.x, xx.x)		(xx.x, xx.x)	(xx.x, xx.x)		(xx.x, xx.x)	
P-value [2]		0.xxxx		0.xxxx	0.xxxx		0.xxxx	

Comparison Between Treatments	Difference [2,3]	95% CI	P-value
Micellar cleanser vs Baxter Sterile Water	xx.x	(xx.x, xx.x)	0.xxxx
Imperial leather Original bar Soap vs. Baxter Sterile Water			
Imperial leather Original bar Soap vs. Micellar cleanser			
Baxter Sterile Water vs.Un-washed			

Supportive Non-Parametric Analysis [4]	(xx.x, xx.x)	0.xxxx

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.

[2] From ANOVA model with change from baseline in examiner assessment on dryness as response and treatment, site as fixed effects. If one-sided upper 95% CI is less than 0.25 then non-inferiority is established.

[3] Difference is first named treatment minus second named treatment such that a negative difference favours first named treatment.

[4] P-value from Wilcoxon sign rank test, 95% CI is constructed on the Hodges-Lehmann method.

PPD

Programming Note:

- Repeat for day 3, 4 and 5, 3 hour post last wash
- Non-parametric method will be presented only if there is violation of normality.
- For the Table 14.2.2.4 the footnotes will be modified as following –

[1] Baseline visit is day 1(AM) visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.

[2] From ANOVA model with change from baseline in examiner assessment on redness as response and treatment, site as fixed effects.

[3] Difference is first named treatment minus second named treatment such that a negative difference favours first named treatment.

[4] P-value from Wilcoxon sign rank test, 95% CI is constructed on the Hodges-Lehmann method.

- For the Table 14.2.2.5 and 6 the footnotes will be modified as following –

- [1] Baseline visit is day 1 visit 2 assessment. Baseline only includes those subjects who have a corresponding post-baseline assessment.
- [2] From ANCOVA model with change from baseline in examiner assessment on redness as response and treatment, site as fixed effects and examiner assessment on TEWL/Corneometer as covariates.
- [3] Difference is first named treatment minus second named treatment such that a negative difference favours first named treatment.
- [4] P-value from Wilcoxon sign rank test, 95% CI is constructed on the Hodges-Lehmann method.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.3.1.1
Summary of Treatment Emergent Adverse Event
Safety Population

Safety Population: N=xx

System Organ Class Preferred Term	Micellar cleanser (N = xx)		Imperial Leather Original Bar Soap (N = xx)		Baxter Sterile Water (N = xx)		Unwashed (N=xx)	
	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE
NUMBER OF SUBJECTS WITH AT LEAST ONE AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NUMBER OF SUBJECTS WITH NO AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN RELATED AES	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN AND SUBCUTANEOUS TISSUE DISORDERS	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
ERYTHEMA	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
DERMATITIS	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NON SKIN RELATED AES	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
GASTROINTESTINAL SYSTEM	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
ABDOMINAL PAIN	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
DRY MOUTH	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
VOMITTING	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx

n (%) = Number (percent) of subjects nAE = Number of adverse events.

Protocol: 207619 Program Run Date:ddmonyyyy

Table 14.3.1.2
Summary of Treatment Emergent Adverse Event by Severity and Treatment/Test Site
Safety Population

Safety Population: N=xx

System Organ Class Preferred Term	Micellar cleanser (N = xx)						Imperial Leather Original Bar Soap (N = xx)						Baxter Sterile Water (N = xx)						Unwashed (N=xx)					
	Mild		Moderate		Severe		Mild		Moderate		Severe		Mild		Moderate		Severe		Mild		Moderate		Severe	
	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE	n (%)	nAE
NUMBER OF SUBJECTS WITH AT LEAST ONE AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
NUMBER OF SUBJECTS WITH NO AE	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx	xx (xx.x)	xx
SKIN RELATED AES																								
SKIN AND SUBCUTANEOUS TISSUE DISORDERS																								
ERYTHEMA																								
NON SKIN RELATED AES																								
GASTROINTESTINAL SYSTEM																								
ABDOMINAL PAIN																								

n (%) = Number (percent) of subjects nAE = Number of adverse events.

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.1.7
Randomisation Information
Randomised Population

Subject Number	Age/Sex/Race[1]	Randomization Number	Date of Randomisation	Treatment Randomised
PPD				

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

Page x of y

PPD

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.2
Individual Subjects Protocol Deviations
Randomised Population

Subject Number	Age/Sex/Race[1]	Visit #	Deviation Sequence	Protocol Deviation
PPD				xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.6
Subjects with Dryness, Redness, TEWL and Corneometer Measurement
ITT Population

Subject Number	Age/Sex/Race[1]	Phase/Visit #	Treatment	Dryness Rating[2]	Redness Rating[3]	TEWL measurement ($g/m^2/hr$)[4]	Corneometer (i.u.)[5]
PPD							xxxxxxxxxxxxxxxxxxxxx

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.

[2] Dryness rating: 0= No dryness; 1.0=Patches of slight powederiness and occasional patches of small scales may be seen; 2.0= Generalised slight powederiness; minimal popular response; 3.0= Generalised moderate powederiness and/or heavy cracking and lifting scales; 4.0=Generalised heavy powederiness and/or heavy cracking and lifting scales; 5.0=Generalised high cracking and lifting scales; 6.0=Generalised severe cracking.

[3]Redness rating: 0=No redness; 1.0=Barely detectable redness; 2.0=Slight redness; 3.0=Moderate redness; 4.0=Heavy or substantial redness; 5.0=Extreme redness; 6=Severe redness.

[4] g=gram; m^2=square meter; hr=hour

[5] i.u.= instrumental units

Protocol: 207619 Program Run Date:ddmonyyyy

Listing 16.2.7
Listing of All Adverse Event
Randomised Population

Treatment Group: Micellar cleanser

Subject Number	Age/Sex/Race[1]	Site	Adverse Event (Preferred Term) (System Organ Class)	Start Date /Study Day[2]	Start Time	End Date	End Time	Frequency /Intensity[3]	Related to Study Product?	Action Taken re Study Product	Outcome	Serious?	Withdrew?[4]
PPD													

[1] Age in years; Sex: F = Female, M = Male; Race: A = Asian, B = Black or African American, I = American Indian or Alaska Native, H = Native Hawaiian or Other Pacific Islander, W = White, O = Multiple.
[2] Study day is the day relative to start of treatment, day 1 being the day of first treatment.
[3] INT = Intermittent and SGLE = Single.

[4] Did subject withdraw from study as a result of this adverse event?

Page x of y

PPD

Programming Note:

- Repeat the same layout for the listing 16.2.7.2
- Population should be used 'Non randomised Subjects'
- The fourth column should be only 'Start Date'
- Delete the footnote related to study day and adjust the numbers accordingly.