Alliance Pharmaceuticals Ltd.

Study VAM2301

Prevention of infestation with *Pediculus humanus capitis* in the family group using a pediculicidal shampoo: A prospective, randomised controlled study

Statistical analysis plan

Author: Dr J S Fry
Date: 20 May 2021
Version: 1.0

QA'd by:	Jan Hamling	Date QA'd
Authorised by:	John Fry	Date authorised
	Jonathan Ormerod (for Alliance Pharma)	Date authorised
	Ian Burgess (for Medical Entomology Centre)	Date authorised

Revision history

Version	Date	Details
1.0	2- May 2021	New

Contents

1 Outline

1.1 Primary endpoints and secondary measures

The primary outcome of the study is the rate of individuals within each family with active head lice infestation at day 14. These rates will be compared between the two treatment groups.

Secondary measures include:

1. Safety of the study products, as determined by the occurrence of adverse events.

2. Subject assessment of acceptability.

The definitions of the primary endpoint and secondary measures are described more fully in the treatment outcome algorithm in section 3.1.2 of the protocol, and in sections 2.9 and 2.12 below.

As the unit of randomisation is the family, with all members of a family given the same treatment, analyses will generally be conducted based on an endpoint evaluated for each family. Thus, for infestation at day 14, comparison will be made between treatments of the mean success rate per family (the primary endpoint) and of the proportion of families where all members satisfied the criteria for lack of infestation, i.e. were successes.

1.2 Outline of the report content

The data processing and statistical analysis of VAM2301 will be presented in a report from RoeLee Statistics Ltd (RoeLee). Apart from describing the methods used and presenting the results of the analyses relating to the primary endpoint and secondary measures, the report will contain details of the study design and objectives, the data items and family and subject numbers available for statistical analysis, and the methods used for data processing. The report will also include results from a variety of additional analyses. For example, comparisons of treatments will be presented for characteristics measured at baseline, the numbers of lice present at the different times of measurement, and the quantity of product used. Although the main results relate to the intention-to-treat (ITT) population, results may also be presented for the per-protocol (PP) population, and for some endpoints, for defined subsets of the population. Analyses showing how the primary endpoint varies by key cofactors (demographic variables, level of initial infestation and hair characteristics) will also be included.

1.3 Structure of the report

The report will start with a cover page giving the name and date of the report and the contact details for RoeLee. Following this will be an executive summary and an index, and then the body of the report, divided into text, tables, listings and appendices, as described below. It is envisaged that the report will not only be complete on its own, but can also be used as an Appendix in the report to be produced by the Chief Investigator.

2 Content of the report

Successive subsections of section 2 refer to the successive sections of the report.

2.1 Introduction

The study design and the objectives of VAM2301 will be described.

2.2 Data provided for statistical analysis

The source of the data will be described and the types of data item available for statistical analysis will be summarised briefly.

2.3 Inclusion and exclusion criteria and populations

The inclusion and exclusion criteria will be listed. Mention will be made of any subjects found to be protocol violators, or terminated prematurely due to adverse events, treatment failure, loss to follow-up or other reasons, as these may have an effect on the subsequent populations to be analysed. As described below, analyses will be conducted on the ITT (intention to treat) population, which includes all participants from randomised families who were treated at least once, and also, if appropriate (see section 2.5), on the PP (per protocol) population, which includes all families in which each participant was treated according to the study protocol. Note that for the purposes of the PP analysis, a study participant will be deemed to be non-compliant if the shampoo is not applied for at least seven (7) days during the study period and/or is not applied for four (4) consecutive days.

2.4 Data processing

The procedures and systems used for accessing the data and creating the database for statistical analysis (a RoeLee database) will be summarised. A list of the subject number, sex, group, family ID and primary outcome (infestation at day 14) will be printed from the original system and the RoeLee database and checked. It is expected that this will match absolutely. There will be an Appendix, which shows the detailed structure of the RoeLee database.

2.5 Statistical methods

Unless the ITT and PP populations turn out to be the same, or differing by only one or two families, separate analyses will be presented for both populations. For some variables, as described further below, ITT analyses will also be conducted for other subsets of the population.

The variables to be analysed are recorded on each subject as either presence/absence variables (e.g. lice present), categorised variables (e.g. infestation level) or continuous or semi-continuous variables (e.g. age or number of lice). As the randomisation is over family, derived variables will be created from the values for each household member within a family to summarise the results for that family. Details of the derived variables used are given in Appendix A to this SAP, but they are themselves typically presence/absence or continuous/semi-continuous. Thus one might have presence/absence variables to indicate whether all members of the family have no infestation at day 14, or whether any member of the family has a given medical condition. Or one might have continuous/semi-continuous variables to indicate the proportion of family members infested at day 14, the mean age of members of a family, or the proportion of family members with long hair ("ears to shoulders" or "below shoulders").

For presence/absence derived variables, exact tests will compare the frequency in families given Vamousse shampoo and in families given the control shampoo. The output will show for each treatment and overall,

N	Number of families analysed
n	Number of families with the condition of interest
%	Percentage of families with condition analysed (100 n/N)
Ex P	Coded exact p value for treatment comparison

For continuous or semi-continuous derived variables, rank tests (Kruskal-Wallis) will be used to compare families given Vamousse shampoo and families given the control shampoo. The output will show for each treatment and overall,

N	Number of families analysed
Mean	Arithmetic mean
Median	Median value

Min	Minimum value
Max	Maximum value
Q25	25th percentile
Q75	75th percentile
p	Coded p value for treatment comparisons

(Note that the mean is presented for information only, and is to be interpreted with caution, as, for some variables, e.g. numbers of lice, the distribution is very non-normal.)

For all statistical tests, two-sided probability values will be coded as follows:

+++, ---	$p<0.001$
++, --	$p<0.01$
+, -	$p<0.05$
(+), (-)	$p<0.1$
N.S.	$p\geq0.1$

Codes will appear in the column related to the second treatment group, with plus signs indicating a greater response in that treatment than in the first treatment group, and minus signs the reverse.

For the primary endpoint some additional analyses will be conducted showing how, within each treatment, the response of subjects varies by the cofactors sex, age, household size, initial infestation level and hair characteristics (hair length, hair thickness, degree of curl and hair type). These analyses will be stratified by family. These analyses will compare success rates for the different levels of each cofactor using stratified chisquared tests, and where the cofactor is a qualitative categorical variable will also include tests for trend to determine whether the success rate increases (or decreases) linearly with successive levels of the cofactor (coded 1, 2, 3 or 4).

Where a cofactor shows a significant ($p<0.05$) relationship with outcome, and where there is evidence of failure of randomisation for that cofactor (i.e. a significant difference between treatments), analyses of the outcome will be

repeated with adjustment for the derived variable for that cofactor. Adjustment will be carried out using stratified chisquared tests (for presence-absence variables) or stratified rank tests (for continuous or semi-continuous variables).

In any event, the main analysis of the primary endpoint will be stratified by family size (represented as <3 children vs ≥ children) and initial infestation level (overt infestations on Day 0=1 versus >1).

While the main analyses compare families given Vamousse shampoo and families given the control shampoo, some additional analyses of adverse event frequency and of frequency of changes in medication will also be carried out comparing individuals given Vamousse shampoo and individuals given the control shampoo, using similar methodology.

Software

The ROELEE system will be used to conduct the analyses. Further details of the statistical techniques used are available from the ROELEE documentation that appears on the RoeLee website (http:/www.roelee.co.uk/Software.htm).

2.6 Tables and listings

This section of the report lists those tables and individual data subject listings that are presented after the text of the report (see also section 3).

The sections that follow (sections 2.7 to 2.13) describe in words various features of the data, each cross-referring to relevant tables and listings.

2.7 Comparability of treated groups at baseline

This section will:

(i) confirm that appropriate consent was obtained for all subjects, and present the frequency of the different types of consent;

(ii) confirm that all the subjects satisfied the inclusion and exclusion criteria;

(iii) compare the groups by age, sex and household size;

(iv) compare the groups as regards hair characteristics (length, thickness, curl and hair type),

(v) compare the groups as regards the extent of the initial infestation,

(vi) compare the groups as regards the previous head lice treatments used and the outcome of the last treatment used, and

(vii) compare the groups as regards each of the 14 broad areas of medical history recorded, and also the more commonly recorded specific conditions.

Group comparisons for (iii) to (viii) will be based on the derived variables for the family.

2.8 Product research

This section will compare the groups as regards the derived variables based on the data recorded on the participant questionnaire (scalp feel, level of odour, acceptability of odour, ease of washing out, feel to touch when dry, ease of brushing or combing following treatment, and overall acceptability).

Counts and listings of these variables will be presented without any formal statistical analysis. These results relate to secondary measure 2 (see section 1.1).

2.9 Completion and withdrawal

This section will summarise the completion rates and reasons for withdrawal.

2.10 Adverse events

This section will describe the adverse events that have occurred. The numbers of subjects and of families having adverse events will be made clear and the distribution of numbers of adverse events per subject will be given, and for frequency, seriousness, severity, action taken and relationship to study treatment, the distribution of worst cases per subject and per family will be summarised. A table will list each adverse event together with its time of onset from treatment, severity, action taken, relationship to study product and MedDRA code details. A shorter form of this table will present data restricted to adverse events with a relationship to study product that is probable or possible. A further table will summarise the numbers of events and numbers of subjects with events by body system.

These analyses relate to secondary measure 1 (see section 1.1).

2.11 Medication

Reference will be made to the actual listings showing the medications current at entry and the changes in concomitant medication given in the listings, but only the proportion of subjects, and of families with subjects, reporting any change in medication since the start of the study will be presented.

2.12 Presence and numbers of lice

For day 3 the groups will be compared as regards the presence of any lice, >5 lice or any small nymphs in the index case. At day 14 the groups will be compared as regards the derived variables for presence of live lice, and for numbers of lice removed, both in total and by type (adult male, adult female, stage 3 nymphs, stage 2 nymphs, stage 1 nymphs).

Note that checking for lice in the index case is only performed at day 6 where lice have been found at day 3, thus no formal analysis can be undertaken at this time point.

2.13 Outcome

The outcome for a subject will be categorised into two groups: 1 = success, that is no lice present at day 14, 2 = failure, that is lice present at day 14. Where the value is unknown due to any reason, such as withdrawal or drop-out, the outcome will be assigned the value of 2: failure. Analyses will be conducted based on the proportion of successes in the family and on whether or not every family member was successful.

This proportion of successes is the primary endpoint.

The principal analysis will be a stratified Kruskal-Wallis test using the Fry-Lee Test. Significance will be declared at a level of 5%, using a two-sided test.

The following two stratifications will be used:

1. <3 children versus ≥3 children
2. overt infestations on Day 0=1 versus >1.

The analyses will present the distribution of this outcome variable over subjects, and also the distribution of the derived values over family (see Appendix A), followed by the results of the exact and chisquared tests for the outcome. For the

derived variable indicating total success in the family, the output will also contain the following additional output:

RR	Relative success rate
RR_ℓ, RR_u	Lower and upper 95% confidence limits of the relative success rate.

Analyses will also be presented testing for variation in the outcomes by sex, age and hair characteristics.

2.14 Product usage

Counts and listings of product usage will be presented. No formal statistical analysis will be undertaken.

2.15 Summary

Section 14 of the report, repeated as the Executive Summary, will summarise the study design and aims, refer to subjects who were protocol violators or did not complete the study, summarise the main characteristics of the populations studied, and give the main results relating to the primary endpoint and secondary measures.

3 Tables, listings and appendices

Following the text of the report will be 10 tables, as follows:

Table 1	Demographic, family and residence data
Table 2	Hair characteristics, infestation assessment and last treatment
Table 3	Medical history
Table 4	Adverse events
Table 5	Changes in medication since day 0
Table 6	Acceptability
Table 7	Presence of lice
Table 8	Outcome
Table 9	Effect of other factors on rate of cure or re-infestation
Table 10	Product usage

Following the tables will be 17 listings, as follows:

Listing 1	Demographics and hair characteristics
Listing 2	Screening data
Listing 3	Informed consent
Listing 4	Dates of first and second treatments and completion/withdrawal
Listing 5	Head lice infestation and previous treatments
Listing 6	Medical history
Listing 7	Participant questionnaire
Listing 8	Inclusion and exclusion criteria
Listing 9	Planned and actual days of assessment
Listing 10	Lice at day 3 for index case
Listing 11	Lice at day 14
Listing 12	Other data days 3, 6 and 14
Listing 13	Description of adverse events and serious adverse advents
Listing 14	Medication and changes
Listing 15	Completion and withdrawal
Listing 16	Outcome
Listing 17	Product usage

The listings will include all the data from the CRF and other relevant material (Adverse Events, Serious Adverse Events, Assessment Summary, Treatment Weights, Family Membership and MedDRA codes), except that information reported at the top of the page will only be shown once, and signature details (already checked to be valid) will not be included. For information recorded on each subject within a family, the data for each variable in the listing will consist of a row for each family sorted on household number within treatment. The columns will give the coded treatment, the family number, the value of the variable for each successive participant within the family (01, 02, 03 …) and the values of derived variable(s) for the family. Data recorded on some subjects but not others, e.g. adverse events, will be in the form of a report for each

relevant participant, giving the complete data recorded relating to the subject of the listing for that participant.

Generally the tables give the distribution and summary statistics for the data shown on an individual subject and family basis in the listings. Exceptionally, some information (including most of the screening data on page 2 of the CRF; details of informed consent; dates of start, treatment, assessment and completion/withdrawal; medical history descriptions; proof of adherence to inclusion and exclusion criteria; comments made in participant questionnaire; full descriptions of adverse events, serious adverse events and concomitant medications; details of usage of individual bottles; and individual treatment times) is shown only in the listings, with no distributions or summary statistics shown.

Following the listings will be Appendix A, Detailed structure of the database.

APPENDIX A

Derived variables to be analysed for each family

Screening data (CRF p2)

For each subject in the family the data will be the same, so the derived variable can be taken to be the data for the participant 01.

The data analysed will be the number of people (a) in the house, (b) checked, (c) with lice, and (d) enrolled in the study.

Identification (CRF p7)

For age the derived variables will be the mean age and the proportion of adults (age 16+).

For sex the derived variable will be the proportion of males.

Hair characteristics (CRF p7)

The derived variables will be:

Hair length:		Proportion of participants with hair "ears to shoulders" or "below shoulders"
Hair thickness:		Proportion of participants with hair "thick"
Degree of curl:		Proportion of participants with hair "wavy" or "curly"
Hair type:	(a)	Proportion of participants with hair "greasy", and
	(b)	Proportion of participants with hair "dry"

Infestation level (CRF p7)

The infestation level for the family will be taken as that of the index case, scoring "light" as 1, "moderate" as 2 and "heavy" as 3.

Previous head lice history (CRF p7)

The derived variable will be whether or not any participant had previously had head lice.

Medical history (CRF p8)

For each of the 14 medical history variables, the derived variable will be whether any member of the family had that history.

Adverse events (CRF pp9-13)

The derived variables will be whether any member of the family had had:

(i) an adverse event,

(ii) a serious adverse event,

(iii) a moderate or severe adverse event,

(iv) an adverse event for which action was taken, and

(v) an adverse event considered possibly or probably related to treatment.

Changes in medication (CRF pp9-13)

The derived variable will be whether any member of the family had had a change in medication.

Completion/withdrawal (CRF p14)

The derived variables will be whether any member of the family

(i) did not complete the study,

(ii) withdrew due to non-compliance,

(iii) withdrew due to loss to follow-up,

(iv) withdrew due to drop out,

(v) withdrew due to lack of efficacy, and

(vi) withdrew due to adverse event.

Numbers of lice (CRF p13)

For day 14 the derived variables will be:

(i) whether any household member has lice present,

(ii) the mean total number of lice per member, and

(iii) the mean number of lice per member of each of the five types (stage 1 nymphs, stage 2 nymphs, stage 3 nymphs, adult males, adult females).

Outcomes (CRF 13)

The derived variables will be the proportion of successes in the family for:

no infestation (no lice present on day 14) – the primary endpoint,

A further derived variables will indicate whether or not every family member was successful.