PPOIT-003LT	Statistical Analysis Plan For Interim Reports
Version 1.0	

PPOIT-003LT Interim Analysis F1

Follow-on study of a multicentre, randomised, controlled trial (PPOIT-003) evaluating the long-term safety and efficacy of Probiotic and Peanut Oral Immunotherapy (PPOIT) compared with Oral Immunotherapy (OIT) alone and with placebo in peanut allergic individuals (PPOIT-003 Follow-On)

Document Version History

Version Date	Version	Author	Signature	Change Description	Reason/Comment
10/5/2022	1.0	Xiaofang Wang		Initial release.	Not applicable.

(FORM: SAP template interim analysis/Version 1.1)

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	Adverse Event
CRF	Case Report Form
DWSU	Desensitization without Sustained Unresponsiveness
F0	First visit of the follow-on study (12 months post treatment)
F1	12 months after F0
F2	24 months after F0
F3	36 months after F0
F4	48 months after F0
F5	60 months after F0
FAIM	Food Allergy Independent Measure
GCP	Good Clinical Practice
GLM	Generalized Linear Model
ITT	Intent-To-Treat
OIT	Oral Immunotherapy
QoL	Quality of Life
SAE	Serious Adverse Event
SD	Standard Deviation
SE	Standard Error
sIgE	Allergen-specific Immunoglobulin E
sIgG4	Allergen-specific Immunoglobulin G_4
SPT	Skin Prick test
SU	Sustained Unresponsiveness
MedDRA	Medical Dictionary for Regulatory Activities
WHODD	World Health Organization Drug Dictionary

1. OBJECTIVES OF THE INTERIM ANALYSIS F1

1. To measure and compare peanut intake and reactions to peanut (accidental or intentional ingestion) from end-of-treatment (T1 of PPOIT-003 study) to 24 months (F1) post-treatment:
   - in PPOIT vs OIT vs placebo and
   - in subjects with SU vs DWSU vs allergic.

2. To compare change from baseline in quality of life (PPOIT-003 study entry) at 24 months (F1) post-treatment:
   - in PPOIT vs OIT vs placebo groups and
   - in subjects with SU vs DWSU vs allergic.

2. BACKGROUND/INTRODUCTION

2.1. GENERAL STUDY INFORMATION

(refer to protocol version 4 dated 05 Oct 2021)

Study design

This is a multicentre follow-on study. There is no treatment intervention in this study. However, there will be study assessments performed including skin prick test (SPT) and blood tests. Subjects will remain blinded to their PPOIT-003 treatment allocation and will continue with peanut avoidance / ingestion according to the instructions they received at the end of treatment in PPOIT-003 as outlined below:

- Subjects who were classified as SU at T2 in PPOIT-003 (Group 1) were instructed to introduce peanut into their diet ad libitum without specific instructions on how often or how much peanut to ingest.
- Subjects who were classified as desensitisation without SU at T2 in PPOIT-003 (Group 2) were instructed to ingest 1 – 2 peanuts each day to maintain desensitisation.
- Subjects who were classified as allergic at T1 in PPOIT-003 (Group 3) were instructed to continue with strict peanut avoidance.

The classification of these clinical outcome groups was determined as follows:

- SU = Subjects who passed both T1 and T2 challenges in PPOIT003
- DWSU = Subjects who were fully desensitised to cumulative 4950mg peanut protein at end of treatment (passed T1 challenge) but failed the T2 challenge
- Allergic = Subjects who failed the T1 challenge irrespective of whether the reaction threshold in T1 challenge had increased from Baseline T0 challenge.

These clinical outcome groups are comprised of subjects in PPOIT-003 who received one of the treatment allocations outlined below however study staff and subjects remain blinded to their treatment allocation:

- PPOIT = Probiotic and peanut OIT taken daily for 18 months during PPOIT-003.
- OIT = Probiotic placebo and peanut OIT taken daily for 18 months during PPOIT-003.
- Placebo = Probiotic placebo and OIT placebo taken daily for 18 months during PPOIT-003.

All subjects will be monitored for 60 months (from 12 months post treatment -F0- to 72 months post treatment -F5).

The study consists of:

Monitoring phase (up to 60 months)

This occurs after completion of the T3 visit of PPOIT-003, which coincides with the first visit of the follow-on study (F0). All subjects will be followed up for 60 months after T3/F0 (which is up to 6 years post treatment that was received in PPOIT-003). During this time, all subjects will be monitored for information on exposure to peanut/amount of peanut being eaten and allergic reactions via a mobile diary application (app) until the end of study (72 months post treatment (F5)).

A phone call will be administered at the following time points:
- Time point F1 – 12 months after F0 (24 months post-treatment)

A visit for SPT, bloods and stool will be administered at the following time points:
- Time point F2 – 24 months after F0 (36 months post-treatment)
- Time point F3 – 36 months after F0 (48 months post-treatment)
- Time point F4 – 48 months after F0 (60 months post-treatment)
- Time point F5 – 60 months after F0 (72 months post-treatment)

Note: Where a face-to-face visit is not possible (eg due to COVID-19 restrictions or other unavoidable circumstances), the visit will be conducted by phone and SPT, bloods and stool test will not be performed.

This study is expected to run for 78 months from the start of participant enrolment to the last participant finishing the study. The length of the study (i.e. follow up period) for each participant is 60 months.

The three sites will be: The Royal Children's Hospital Melbourne / Murdoch Children's Research Institute, The Women's and Children's Hospital Adelaide, Perth Children's Hospital. A maximum of 201 participants will be recruited – the study population recruited to PPOIT-003.

Study population
Subjects meeting the eligibility criteria will be recruited at end of PPOIT-003 without unblinding of PPOIT-003 treatment allocation.
Inclusion criteria
Subjects are eligible for the study if they meet all of the following criteria:
- Were randomised and received at least 1 dose of treatment in PPOIT-003
- Participated in the PPOIT-003 randomised trial
- Written informed consent from participants and/or parent/guardian.

Exclusion criteria
Subjects are not eligible for the study if they meet any of the following criteria:
- Are taking probiotic supplements (does not include formula, yoghurts and fermented foods).
- Withdrew from the PPOIT-003 study.
- Have any condition that, in the opinion of the Investigator, precludes participation for reasons of safety.

Sample size

No formal sample size calculation is being conducted for the follow-up study. A maximum of two hundred and one (201) participants will be recruited from PPOIT-003 trial (PPOIT n=79; OIT n=83; placebo n=39).

2.2. DESCRIPTION AND SCHEDULE OF INTERIM ANALYSIS

Interim analyses of the F1, F2, F3 and F4 data will be performed after all enrolled subjects have completed the F1, F2 and F3 study time points or withdrawn from the study at F1, F2 and F3 study timepoints, respectively. Analysis will be conducted using available data collected to the F1, F2 and F3 time point and will comprise primary and secondary outcomes (where relevant). This SAP is for the interim analysis of the F1 data. Summary tables will report data aggregated by treatment group and/or clinical outcome, with no individual subjects being identified.

Unblinding
Subjects' treatment allocation from PPOIT-003 will remain blinded throughout the duration of this follow-on study. Subjects will keep the same study number allocated from the parent study. Trial drug codes of the PPOIT-003 trial will be made available only to the unblinded study statistician and other unblinded independent consultants after database lock at the interim analysis and at the end of the study. All other study staff who are involved in the PPOIT-003 follow on study and assessments on the subjects, as well as the subjects themselves, will remain blinded to the allocation received during the PPOIT-003 study until the completion of the follow-on study.

3. POPULATIONS OF INTERIM ANALYSIS

The analysis of all outcome data will include all subjects who were randomised to PPOIT-003 trial and consented to roll over to PPOIT-003LT, where outcome data is available at the F1 time point.

4. VARIABLES

4.1. OUTCOMES

Table 1 –Outcomes description

	Outcome	Description
A	**Peanut ingestion (intentional or accidental) between T1 and F1**	Binary outcome 1=Yes if ingested peanut at least once over the 12 months period. 0=No
B	**Reaction to peanut ingestion (intentional or accidental) between T1 and F1**	Binary outcome 1=Yes if reacted to peanut at least once over the 12 months period 0=No
C	**Number of reactions to peanut ingestion (intentional or accidental) T1 and F1 <u>overall</u> and <u>by severity</u>**	Continuous outcome and Categorical outcome 1=1 reaction 2=2 reactions 3=3 or more reactions
D	**Use of EpiPen to treat reaction to peanut (intentional or accidental) between T1 and F1**	Binary outcome 1=Yes if used EpiPen to treat reaction to peanut at least once over the 12 months period 0=No
E	**Number of times using EpiPen to treat reactions to peanut (intentional or accidental) between T1 and F1**	Continuous outcome
F	**Weekly average quantity of peanut protein ingested (intentional or accidental) between F0 and F1**	Continuous outcome and Categorical outcome 1=Never ingested 2=Minimal amount (less than 1 peanut, <250g) 3=Small amount (1-2 peanuts, 250-599g) 4=Moderate amount (3-8 peanuts, 600-1999g) 5=Large amount (>2000g)
G	**Weekly average frequency of peanut ingestion (intentional or accidental) between F0 and F1**	Continuous outcome and Categorical outcome 1=Nil 2=Less than once a month 3=Once a month to <1/week 4=1-2 times/week 5=3-5 times/week 6=>5 times a week
H	**Quality of life scores – total, sub-scores (emotional impact, food anxiety, social and dietary limitations) and FAIM total score at F1**	Continuous outcome. From FAQLQ-PF, total score is calculated as the mean of up to 30 items (some are age specific), with response scale from 0 (minimal impairment in QoL) to 6 (maximal impairment in QoL).

		The mean scores for selected items are used to construct the sub-scores. The FAIM consists of 4 questions, scored on a 7-point scale. The FAIM total score is the mean of all items and ranges from 0 (limited severity perception) to 6 (greatest severity perception).
I	**Change in quality of life scores (total, sub-scores, and FAIM) from screening in PPOIT-003 study to F1**	Continuous outcome

4.2. OTHER VARIABLES

These variables are for purposes of describing the study sample.

Characteristics at the screening visit of PPOIT003: age, weight/height, medical history, skin prick test (peanut, egg, milk, cashew, almond, pistachio, hazelnut, dust mite and positive and negative control), blood and faecal sample (faecal sample collected by parents at home), anaphylaxis education.

Characteristics at the F0/T3 visit: age, weight/height, medical history, skin prick test (peanut, egg, milk, cashew, almond, pistachio, hazelnut, dust mite and positive and negative control), blood and faecal sample (faecal sample collected by parents at home), anaphylaxis education.

Characteristics at the F1 time point:

- *An Allergy Questionnaire* collected information on allergic disease (food allergies, asthma, eczema), medications, and peanut ingestion (accidental or intentional, amount, frequency).

- Parents of children aged 0-12 years will complete a validated *Food Allergy Quality of Life-Parent Form* (FAQLQ-PF) to report their child's health-related quality of life from the child's perspective. Children aged 8–12 years will complete the *Food Allergy Quality of Life-Child Form* (FAQLQ-CF) to specifically capture quality of life from the child's perspective. Children aged 13–17 years will complete the *Food Allergy Quality of Life-Teenager Form* (FAQLQ-TF) to specifically capture quality of life from the teenage child's perspective.

5. STATISTICAL METHODOLOGY

5.1. F1 GENERAL METHODOLOGY

Outcomes will be analysed by treatment group, according to the group to which subjects were originally randomly allocated, regardless of subjects' compliance to treatment during the PPOIT-003 study, or crossover to other treatments. The following pairwise comparisons will be of interest:

- PPOIT vs Placebo
- PPOIT vs OIT
- OIT vs Placebo

Outcomes will also be described by the clinical outcome achieved at end of treatment in PPOIT-003 (T2 PPOIT-003), that is SU, DWSU without SU and Allergic. The following pairwise comparisons will be of interest:

- SU vs DWSU
- SU vs Allergic
- DWSU vs Allergic

All demographic and continuous outcomes (collected at T0 and T3/F0) will be presented as mean and standard deviation (or medians and interquartile ranges for skewed data), whilst categorical outcomes will be presented as absolute and relative frequencies in the three treatment groups and clinical groups. Table of characteristics at T0 will be presented by treatment groups for those consented to roll over to PPOIT-003LT; Table of characteristics at T3/F0 will be presented both by treatment groups and by clinical outcome groups.

In this interim analysis, subjects who are withdrawn at the end of the PPOIT003 study will not be included in any analyses. All the outcomes will be reported in available cases who consented to roll over to PPOIT-003LT.

MEASURES OF TREATMENT EFFECT
Risk difference (RD) and relative risks (RR) between treatment groups (or clinical outcome groups) will be the measures of treatment effect for binary outcomes.
Mean difference between treatment groups (or clinical outcome groups) will be the measure of treatment effect for continuous outcomes whose distribution is reasonably symmetrical.
Median difference between treatment groups (or clinical outcome groups) will be the measure of treatment effect for continuous outcomes whose distribution is severely skewed.

ESTIMATION OF TREATMENT EFFECTS (BETWEEN TREATMENT GROUPS)
For binary outcomes, a generalized linear model (GLM) approach will be used and account for clustering by site, adjusted by the age group at randomisation (1-5 years; 6-10 years) and SPT wheal size (≤10mm; >10mm).

For continuous outcomes, the mean difference between treatment groups will be obtained by using GLM to account for clustering by site, adjusted by the age group at randomisation (1-5 years; 6-10 years) and SPT wheal size (≤10mm; >10mm). The median difference using quantile regression model to adjust for the stratification factors if not normally distributed.

We will also compare the QoL scores using a mixed model applied to the data at 4 timepoints (T0, T1, T3/F0, F1) and account for clustering by study site, with fixed effects for treatment, age group at randomisation and SPT wheal size, and a random effect to allow for repeated measures for each individual.

ESTIMATION OF EFFECTS (BETWEEN CLINICAL OUTCOME GROUPS)
For binary outcomes, a binary regression model will be used.

For continuous outcomes, the mean difference between the clinical outcome groups will be obtained by using linear regression. Adjusted linear regression will be used to adjust for baseline QoL score in outcome I. The median difference using quantile regression model if not normally distributed.

HANDLING OF MISSING DATA
It is anticipated that 25% of children will not continue to participate the follow-up, due to participant attrition and parent/guardian non-response. The available case analysis will be the

primary one. A sensitivity analysis using Inverse Probability Weighting (IPW) and Multiple Imputation (MI) techniques will be used to handle the missing data (details in section below).

SENSITIVITY ANALYSIS

The sensitivity analysis will be performed on an intention-to-treat basis including all the children whose parent/guardian consented for follow-up, with IPW and MI techniques used to handle the missing data. It is anticipated that 25% of the original sample of 201 children will have follow-up outcomes missing. In such situations, we will use both MI combined with IPW techniques (Seaman, S.R., White, I.R., Copas, A.J. and Li, L., 2012. Combining multiple imputation and inverse-probability weighting. Biometrics, 68(1), pp.129-137.) as a sensitivity analysis. It involves two major steps:

Step 1, calculation of Inverse Probability Weights: Inverse probability weights will be calculated by treatment group using the whole sample of 201 children adopting a logistic regression model predicting participation at F1, using the following baseline variables (with no missing values): age at randomisation and SPT wheal size, clinical outcome at T2 (SU, DWSU without SU or Allergic) Once inverse probability weights are obtained, only children whose parent/guardian consented for follow-up will included in the following step, and in the final analyses.

Step 2, multiple imputation of missing data on those who participated in the F1 follow-up: The frequency and patterns of missing data will be examined. Missing values of included participants will be multiply imputed: multiple imputation models will be specified including all variables included in the analysis models (baseline stratification factors) as well as all the outcomes collected at T3/F0, and F1 follow-up. Thirty completed data sets will be imputed using multivariate normal regression including all the participants who participated in the F1 follow-up. Each dataset resulting from the multiple imputation model (a “quasi-complete dataset” because the data are complete for the included participants, but not for those excluded from MI due to not participating in the F1 follow-up) is then analysed using IPW derived at step 1. IPW reweights data so that it accounts for characteristics of excluded participants (those whose parent/guardian did not consent for the follow-up and are missing all outcome data).

By imputing for participants with few missing values, while using inverse probability weights to account for excluded participants with all missing outcome data, IPW/MI should maintain the efficiency advantage of MI while avoiding possible bias from incorrectly imputing larger blocks of data.

All outcomes listed in Table 2 will be analysed using IPW/MI as a sensitivity analysis.

SUB-GROUP ANALYSIS – AGE GROUP AT SCREENING OF PPOIT003

Available cases will be analysed by subgroups according to their age group at screening of PPOIT003 (0-5 years / 6-10 years) for all outcomes.

5.2. F1 INTERIM ANALYSIS METHODOLOGY

Outcomes A, B, D

For these binary outcomes, they will be summarised as the number and proportion of children who had the outcome as yes between F0 and F1 by treatment groups and by clinical outcome groups.

The comparison between treatment groups will be made using a GLM approach with Gaussian error distribution (to avert convergence difficulties with low prevalence outcomes), linear link function and a cluster-robust standard error calculation to account for clustering by study site. The RD and 95% CI will be derived, adjusted by the age group at randomisation (1-5 years; 6-10 years) and SPT wheal size (≤10mm; >10mm).

The RR (with 95% CI) comparing treatments will also be estimated using a GLM approach (the "modified Poisson" approach of Zou, employing the Poisson family, a log link function and a cluster-robust standard error calculation to account for clustering by site), adjusted by the age group at randomisation (1-5 years; 6-10 years) and SPT wheal size (≤10mm; >10mm).

The comparison between clinical outcome groups will be made using a binary regression model. The RD, RR and 95% CIs will be derived.

Outcome C, F, G

For these categorical outcomes, they will be summarised as the number and proportion of children who had the outcome in each category by treatment groups and by clinical outcome groups. They will also be summarised as the median, IQR and range by treatment groups and by clinical outcome groups.

Outcomes C, E, F, G, I

For these continuous outcomes, they will be summarised as mean and SD or median and IQR by treatment groups and by clinical outcome groups as specified.

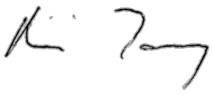
The mean difference between treatment groups will be obtained by using GLM to account for clustering by site, adjusted by the age group at randomisation (1-5 years; 6-10 years) and SPT wheal size (≤10mm; >10mm). The median difference using quantile regression model to adjust for the stratification factors if not normally distributed.

The mean difference between clinical outcome groups will be obtained by using a linear regression model. The median difference using quantile regression model if not normally distributed. For outcome I, linear/quantile regression model will be used for further adjusted for baseline QoL score.

Outcome H

We will also compare the QoL scores between treatment groups using a mixed model applied to the data at 4 timepoints (T0, T1, T3/F0, F1) and account for clustering by study site, with fixed effects for treatment, age group at randomisation and SPT wheal size, and a random effect to allow for repeated measures for each individual.

SIGNATURES PAGE

Signature of Group Leader:		Date 04 -Jun- 2022
Print Name	Prof Mimi Tang	
Signature of Principal Investigator:		Date 31-May-2022
Print Name	Dr Paxton Loke	
Signature of Trial Statistician:		Date 10-May-2022
Print Name	Francesca Orsini	