NCT 02679807

30 July 2018

Effect of Probiotic on Rhinovirus Induced Colds

Study code: MK EPRIC

Phase II study

STATISTICAL ANALYSIS PLAN

Signatures:

Statistical Analysis Plan was prepared by:

Sofia Männikkö Study Statistician	Date

Statistical Analysis Plan was reviewed/approved by:

Teppo Huttunen Senior Statistician	Date
Markus Lehtinen	Date
Liisa Lehtoranta	Date
Ronald Turner	Date

Table of Contents

1 Abbreviations

AE	Adverse event
ANOVA	Analysis of variance
ATC	Anatomical Therapeutic Chemical Classification System, (a WHO drug classification system)
IL-8	Interleukin-8 assay
ITT	Intention-to-treat
MedDRA	Medical dictionary of regulatory authorities
PP	Per-protocol
qPCR	Quantitative polymerase chain reaction
RM-ANOVA	Repeated measures analysis of variance model
SAE	Serious adverse event

2 Study objective(s)

The primary objective of this study is to investigate the effect of probiotic *Bifidobacterium animalis* subsp. *lactis* Bl-04 on incidence of rhinovirus-associated common cold illness episodes in comparison with placebo. The results of this study should permit the decision of whether probiotic supplementation reduces the incidence of rhinovirus induced cold episodes compared to placebo.

The secondary objectives of this study are to investigate the effect of probiotic supplementation on the infection rate and duration of illness episodes, examine the effect of the probiotic on the rhinovirus load in the nose and to study the safety and tolerability of the product.

3 Design and type of the study

This is a phase II randomized, double-blind, placebo-controlled study of probiotic *Bifidobacterium animalis* subsp. *lactis* Bl-04 on incidence of common cold illness episodes induced by rhinovirus in healthy adults.

4 Sample size considerations

A total of 220 subjects were to be challenged and distributed in two groups of equal sizes to the two treatment groups. Based on the sample size calculations 95 completed subjects/ arm would be expected to detect a reduction in rhinovirus-induced illness of 20% (from 60% to 40%) with 80% power with $p_\alpha = 0.05$.

During the study, to ensure reaching the target sample size of completed subjects in the primary analysis population (Per Protocol dataset), the number or challenged subjects was decided to be increased to 254 in total.

5 Statistical hypotheses

The purpose of this study is to support the evidence that the probiotic supplementation is more effective than the effect of placebo in reducing the incidence of rhinovirus-associated illness episodes, i.e., the null-hypothesis to be tested is

H_0 : There is no relationship between the treatment and the incidence of rhinovirus-associated illness episodes

H_1 : There is a relationship between the treatment and the incidence of rhinovirus-associated illness episodes

The effect of probiotic supplementation is considered shown if the two-sided p-value of the primary statistical analysis is less than 0.05.

6 Analysis sets

6.1 Safety dataset

All randomised subjects receiving the study products will be included in the safety dataset.

6.2 Intention-to-treat (ITT) dataset

The ITT dataset will consist of all subjects having been randomized to a treatment (Safety dataset) and receiving the virus challenge on Day 0 per exclusion criteria defined in the protocol. This population will be used for all efficacy analyses (see 20.2.).

6.3 Per protocol (PP) dataset

The PP dataset will consist of all subjects having been randomized to a treatment and receiving the virus challenge on Day 0 (ITT), without any major protocol violation. Subjects with acute neutralizing titer of >1:4, respiratory virus detected prior to virus and antibiotic use before day 14 will be excluded, and only subjects with study product compliance of >80% will be included. The subject classification to PP dataset will be done before opening of the randomisation code. PP dataset is used for all efficacy and exploratory analyses as the primary dataset, thus also for the primary evaluation of study hypothesis (see 20.2.).

6.4 Infected dataset

The Infected dataset will consist of all subjects in the PP dataset, who got infected with the challenge virus. Infection will be defined as the isolation of the challenge virus on any of the 5 days (study days 1-5) after virus challenge or a four-fold rise in neutralizing antibody titer to the study virus. The Infected dataset will be used for analysis of e.g. symptom score, viral load and exploratory analyses (see 20.2).

6.5 Ill dataset

All subjects who have met the criteria for having an illness episode, based on criteria set in assessment of primary outcome (11.1). This dataset is a subset of the PP dataset and will be used to analyse the duration of illness.

7 General statistical considerations

In the statistical analyses, a p-value less than 0.05 will be considered as statistically significant. If not stated otherwise, all tests will be performed as two-sided tests and two-sided 95% confidence intervals will be produced for the treatment differences. No adjustments for the p-values will be made. Missing values will not be imputed in the analyses.

7.1 Examination of subgroups

The following subgroups for the primary endpoint will be investigated: antibody status (four-fold rise in neutralizing antibody titer to the study virus: yes/no), cohort, gender and incidence of wild-type virus.

8 Demographic and other baseline characteristics

The comparability of the treatment groups will be assessed by the demographic and baseline characteristics. These include but are not restricted to

- demographics (age, sex, race, ethnicity)
- effective use of birth control
- habits (physical activity, smoking)
- assessment of blinding
- medical history and concomitant diseases
- physical examination
- vital signs

All baseline variables will be reported using descriptive statistics. If feasible, the baseline distribution will be evaluated with inferential statistics for flagging purposes. One-way ANOVA (or appropriate non-parametric test) between the treatment groups will be applied for continuous variables such as age. For categorical variables chi-square or Fisher's exact test will be used.

9 Prior and concomitant medication/treatment

Concomitant medication and treatments will be summarized by treatment group using ATC classification.

10 Extent of exposure and compliance

Extent of exposure will be evaluated descriptively as duration of study treatment (number of days).

Compliance will be assessed by summarizing the number of used and unused sachets returned as percentages of the total number of sachets dispensed. Product recovery from fecal samples will be evaluated by qPCR.

11 Analysis of efficacy

11.1 Primary efficacy variables

The primary efficacy analysis is comparing the incidence of rhinovirus-associated illness episodes (Yes/No) between the treatment groups. A rhinovirus-associated illness episode is detected in case both the following criteria are met:

1) Rhinovirus infection: Isolation of the challenge virus on any of the 5 days (study days 1-5) after virus challenge OR a four-fold rise in the neutralizing antibody titer to the study virus (day 28).
2) Symptomatic illness: Based on the modified Jackson criteria a total symptom score of at least 6 over the 5 days (sum of study days 1-5) after virus challenge AND either three days of rhinorrhea or the subjective impression that they have had a common cold illness.

The relationship between the treatment and the incidence of rhinovirus-associated illness episodes will be tested by using a chi-square test. In case the cell frequencies, such as the expected number of subjects experiencing an illness episode or the expected number of subjects without any episodes in either of the treatment groups, turn out 5 or lower, a Fisher's exact test will be used for the comparison instead, which is a more suitable method with low expected frequencies.

11.2 Secondary efficacy variables

11.2.1 Duration of illness

Duration of illness is defined as the time between the virus challenge and the end of the illness episode in volunteers who met the modified Jackson definition of symptomatic illness (11.1). The last day of the illness will be defined as the first day of the two consecutive days with a total symptom score ≤ 1 that occurs after the subject has met the symptom criteria for a Jackson cold. In case the symptom score is ≤ 1 on the last day of symptom score assessment, the end of illness will be set to the last assessment day. If the illness continues >14 days (symptom score >1 on the last day of assessment), the length of illness will be set as 15 days.

The main analysis will be done using the Ill dataset. Additionally, analysis will be repeated with PP and ITT datasets. For these additional analyses using datasets including also non-ill subjects, the duration of illness will be set to zero for subjects experiencing no illness.

Illness duration will be analysed by comparing the group mean durations with a two-sample independent t-test. In case the durations are not normally distributed a non-parametric Wilcoxon-Mann-Whitney test will be used for the comparison.

11.2.2 Incidence of infection

Incidence of infection (Yes/No) will be defined as the isolation of the challenge virus on any of the 5 days (study days 1-5) after virus challenge or a four-fold rise in neutralizing antibody titer to the study virus.

Incidence of infection will be analysed exactly the same way as the primary response (section 11.1).

11.2.3 Viral load during infection

Viral load will be assessed by comparing the quantitative virus titer in nasal lavage between the probiotic and placebo groups. For this calculation, subjects with no virus isolate on a study day will be defined as having a quantitative titer of – 0.5 ($\log_{10}$).

The quantitative viral titer load during infection will be evaluated with repeated measures analysis of variance model (RM-ANOVA) including the daily titer load as a response (log_{10}), treatment, day and the interaction of day and treatment as explanatory variables. The random effect of subject will be included in the model. Model assumptions will be checked with model diagnostics and normality plots. In case the assumptions of the RM-ANOVA model are severely violated (after common transformations), generalised linear mixed models for repeated measurements with more appropriate distribution assumptions (e.g. poisson or negative binomial) will be fitted. The main interest of this analysis is the overtime treatment effect.

11.2.4 Proportion shedding the virus

Virus shedding will be defined as the isolation of the challenge virus on any of the 5 days (study days 1-5) after virus challenge.

Proportion shedding the virus will be analysed exactly the same way as the primary response (section 11.1).

11.2.5 Acute illness period total symptom score

Daily total symptom score during study days 1-5 will be analysed based on the Jackson criteria with repeated measurements models similarly than the quantitative viral titer load (11.2.3). The main interest of this analysis is the overtime treatment effect. Individual symptom scores will be evaluated as exploratory endpoints (described in 11.3).

11.2.6 Total symptom score based on the Jackson criteria

Daily total symptom score during study days 1-14, will be analysed based on the Jackson criteria with repeated measurements models (like in 11.2.3). The main interest of this analysis is the overtime treatment effect.

The secondary analyses are ordered in a hierarchical order by their clinical importance as described in this section. Thus, no multiplicity correction is needed, and a p-value ≤ 0.05 can be stated as statistically significant as long as the hierarchy is being taken into account with the conclusions.

11.3 Exploratory analyses

Individual symptom scores based on the Jackson criteria

The individual symptom scores for obstruction, rhinorrhea, cough and sore throat will be analysed similarly as the daily total symptom scores (11.2.5 and 11.2.6).

Cytokine response of nasal mucosa

Change in Interleukin-8 assay (IL-8) concentrations both from day -28 until virus challenge and from day 0 onwards will be analysed with repeated measures analysis of covariance models (RM-ANCOVA), by including the change from day -28/ day 0 as a response, treatment, baseline covariate (IL-8 at day -28/ day 0), timepoint and the interaction of timepoint and treatment as explanatory variables. The random effect of subject will be included in the model. Model assumptions will be checked with model diagnostics and normality plots. In case normality

assumptions cannot be confirmed, common transformations (log, square root, inverse) will be used. For the analysis 'not detected' values will be replaced with the detection limit / 2. For the change from day -28 analysis the PP dataset will be used, and for the change from day 0 analysis the infected dataset will be applied.

Multiplex assays: A panel of cytokines and chemokines

A panel of cytokines and chemokines (the following parameters will be analysed: IFN-y, IL-10, G-CSF, MCP-1, IP-10, IL-6, IL-1b) measured will be analysed in exactly the same way as IL-8.

Rhinovirus type-specific antibody response

The effect of probiotic on host antibody responses of the challenge virus (convalescent titer) will be analysed with one-way ANOVA, by including the change from day 0 to day 28 as a response and treatment as a fixed effect. Model assumptions will be checked with model diagnostics and normality plots. In case normality assumptions cannot be confirmed, common transformations (log, square root, inverse) will be used.

Faecal, throat and nasal microbiota

For the faecal, throat and nasal microbiota at least the following analyses will be conducted. Faecal results will be analysed only descriptively.

For qPCR detecting the supplemented probiotic strain at least frequencies (Positive, Negative) by treatment and time point will be presented. If feasible, proportions can be compared with logistic regression models. For throat and nasal results the model will include the fixed effects of treatment and timepoint, the interaction of treatment and timepoint as a baseline covariate. The random subject effect will be included in the model. Also, a similar model including only observations after day 0 will be fitted to assess the effect of the virus.

For all faecal, throat and nasal microbiota, diversity of the samples will be investigated by comparing the alpha diversity (within sample diversity) values between treatments and timepoints. Alpha diversities will be calculated at DuPont and provided to 4Pharma for further analyses. For throat and nasal microbiota, the change in alpha diversity will be analysed with RM-ANCOVA in exactly the same way as described for IL-8.

From throat swabs, abundance (+, ++, +++) of 5 pathogens of special interest (*S. aureus*, pneumococcus, *H. influenza*, meningococcus and group A streptococcus) will be presented with frequency tables and analysed with cumulative logistic regression model. Probability of higher incidence will be modelled and day zero result and treatment included as fixed effects.

Correlation of abundance of the 5 pathogens with antibody response, incidence of infection or rhinovirus will be investigated with Spearman correlation coefficients.

If feasible, clustering of the samples into microbiota subtypes will be investigated separately for faecal, throat and nasal microbiota based on the weighted UniFrac metric distance matrix including all samples from all time points. UniFrac metric distance matrix will be calculated at DuPont and provided to 4Pharma for the clustering and further analyses. Unweighted pair-group method with arithmetic averages or other feasible method will be used for the clustering. The decision of the optimal number of clusters will be made based on peak values in the pseudo F and pseudo t-squared values or other suitable method. If clusters are formed, difference in proportions by treatment will be investigated with categorical analysis models and the cluster effect on secondary outcome

variables will be explored by adding the cluster effect to the analysis models as a fixed effect. Effect of cohort on clustering will be investigated with categorical repeated measures model, where clustering is the outcome variable, and timepoint, treatment, cohort, interaction of treatment and timepoint are included as fixed effects.

Blood transcriptomics

For blood transcriptomics at least the following analyses will be conducted. The PP population will be primarily used.

Unsupervised clustering of the samples will be conducted. Clustering will be done at DuPont and in case clusters are formed, the clustering information (per sample per subject) will be provided to 4Pharma for further the analyses described below.

- Timepoint and treatment effect on clustering will be investigated with categorical repeated measures models, where clustering is the outcome variable, and timepoint, treatment, interaction of treatment and timepoint are included as fixed effects.
- Effect of cohort, infection (yes/ no) and illness (yes/ no) on clustering will be investigated with the above described model by including separately each factor to the analysis model.
- The effect of D-28/D0 clustering on the following outcome variables will be investigated by adding the cluster as a fixed effect to the analysis models: Duration of illness, viral shedding, clinical symptom severity score and immune markers. For symptom scores, the analysis models have already been described in 11.2.5 and 11.2.6. Also, for immune markers the analysis model to be used has already been described in this chapter. For duration of illness an analysis of covariance (ANCOVA) model will be used including treatment and cluster effect in the model. The analysis will be conducted using the Ill population. For viral shedding a logistic regression model with treatment and cluster effect will be fitted and infected population used. For all models, also fixed interaction effect of treatment and cluster may be investigated.

In addition to the clustering, correlations between individual genes from throat and nasal metagenomics will be calculated, if feasible. For the correlations FDR correction will be used with the p-values.

Connections between the exploratory data/ results and efficacy and other endpoints may be investigated with suitable analysis models (will be described in the Documentation of the Statistical Methods, if done). Part of the exploratory analyses (microbiota and transcriptomics) will not be included in the Statistical Study Report.

11.4 Sensitivity analyses

As a sensitivity analysis, a logistic regression will be computed for the primary efficacy variable. In addition to the treatment effect also potential confounding factors, such as antibody status, cohort, gender or incidence of wild-type virus, are one by one included to the model as fixed effects. Also, the interaction effect (subgroup*treatment) will be investigated.

12 Cold questionnaire comparison

Alternative cold questionnaire from earlier study conducted by West et al. (AUSTRALIA) study will be used to define the illness episodes. Illness episodes defined based on the modified Jackson criteria will be compared with the AUSTRALIA defined episodes and investigated if both questionnaires identify the same episodes. Number of episodes identified by both questionnaires and only one questionnaire will be summarized descriptively. Since training was part of the AUSTRALIA criteria, if the subject had been training was asked also in this study in addition to the modified Jackson criteria. Thus AUSTRALIA defined episodes will be compared both to modified Jackson with training and to modified Jackson regardless of training defined episodes.

13 Analysis of safety and tolerability

All subjects who have been randomised to the study and used the study products will be included in the safety analysis.

13.1 Adverse events

Adverse events will be tabulated by treatment group, system organ class, preferred term, causality and severity. Both subject and event counts will be calculated. In addition, SAEs and AEs leading to discontinuation will be summarized. Adverse events reported for more than 5 % of the patients in either treatment group will be tested using the Fisher's exact test.

14 Completion and premature discontinuation

Completion and premature discontinuation will be listed and the reasons for premature discontinuation will be presented.

15 Deviations from the analyses planned in the study protocol

Secondary analyses of 11.2.5 and 11.2.6 for Acute and Total Symptom Score based on the modified Jackson criteria were added to the hierarchy. Also, individual symptom scores based on the Jackson criteria were added to the exploratory endpoints (11.3).

16 Meta-analysis

A meta-analysis combining the EPRIC and earlier study conducted by Turner et al. (EPIARR) results will be conducted. The primary response for the meta-analysis will be the incidence of infection, other assessments to be investigated include but are not limited to viral load, proportion shedding and IL-8 response. Results of the meta-analysis will be reported separately.

17 Execution of statistical analyses

Statistical analyses will be performed by 4Pharma Ltd.

18 Hardware and software

Statistical analysis, tables and patient data listings will be performed with SAS® version 9.3 or higher for Windows (SAS Institute Inc., Cary, NC, USA).

19 References

Clinical Study Protocol, Final Protocol (17-08-2016), Danisco Sweeteners Oy.

Turner RB, Woodfolk JA, Borish L, Steinke JW, Patrie JT, Muehling LM, Lahtinen S, Lehtinen MJ. Effect of probiotic on innate inflammatory response and viral shedding in experimental rhinovirus infection - a randomised controlled trial. Benef Microbes. 2017 Apr 26;8(2):207-215.

West N, Horn P, Pyne D, Gebski V, Lahtinen S, Fricker P, Cripps A (2014) Probiotic supplementation for respiratory and gastrointestinal illness symptoms in healthy physically active individuals. Clinical nutrition 33: 581-587.

20 Appendices

20.1 Schedule of events

	Screening	Day -28	Day -21	Day -14	Day -7	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6-14	Day 21-28
Consent	X												
Review of inclusion/ exclusion	X					X							
Serum neutralizing antibody	X*					X							X
Urine Pregnancy Screen for females		X				X							
Start of dietary run-in period	X												
Compliance check/ resupply			X	X	X	X							
Probiotic/ placebo treatment		X	X	X	X	X	X	X	X	X	X	X	
Assessment of blinding						X							
Nasal swab for microbiomics		X				X			X				
Throat swab for microbiomics		X				X			X		X	X (14)	
Throat swab for culture						X					X		
Fecal sample for microbiomics						X							
Blood for transriptomics		X				X		X					
Nasal lavage for viral PCR						X							
Nasal lavage for IL-8/ multiplex		X		X		X	X	X	X	X	X		
Symptom assessment						X	X	X	X	X	X		
Rhinovirus challenge						X							
Nasal lavage for quantitative RV							X	X	X	X	X		

culture													
Symptom diary												X	
Assessment of Adverse Events			X	X	X	X	X	X	X	X	X	X	X
Assessment of Concomitant Medications			X	X	X	X	X	X	X	X	X		X

*Volunteers will be screened for neutralizing antibody under an existing screening protocol at the University of Virginia (IRB-HSR #9948). Volunteers who are found to have antibody titers ≤1:4 to the study virus will be invited to consent for participation in the EPRIC study. Only volunteers who were tested more than 12 weeks prior to the screening visit will have serology repeated at this visit.

20.2 Analyses per dataset

The following datasets will be used primarily for the analyses. For exploratory and post-hoc analyses also other datasets may be used.

		Dataset				
	Variable	**ITT**	**PP**	**Inf.**	**Ill**	**Safety**
Primary	Incidence of rhinovirus-associated illness episode	x	**X**			
Secondary	Duration of illness	x	x		**X**	
	Incidence of infection	x	**X**			
	Viral load	x	x	**X**		
	Proportion shedding	x	x	**X**		
	Total symptom score (acute illness)	x	x	**X**		
	Total symptom score (illness)	x	x	**X**		
Exploratory	Each symptom score (eg rhinorrhea; sneezing, obstruction, sore throat)		x	**X**		
	IL-8 + cytokines		**X**	**X**		
	Ab response			**X**		
	Fecal samples - Bl-04 PCR		**X**			
	Nasal swabs		**X**	x		
	Throat swabs microbiomics		**X**	x		

	Throat swabs culture	**X**	x		
	Transcriptomics	**X**	x	X	
Safety parameters					**X**

- **X** = main analysis; x = secondary analysis

20.3 Table and figure plan

See separate excel document.

20.4 Data listing plan

See separate excel document.