

Statistical Analysis Plan

RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, STUDY OF THE EFFICACY OF PHENYLEPHRINE HCL [REDACTED] 30 MG AND PHENYLEPHRINE HCL [REDACTED] 12 MG CAPSULES IN SUBJECTS WITH NASAL CONGESTION DUE TO THE COMMON COLD

Investigational Product Name: Phenylephrine HCl [REDACTED] [REDACTED] ablet, 30 mg

Protocol Number: CO-170302131230-URCT

IND / EudraCT number: IND: 121517

Phase: 2

Version: Date FINAL: 31 AUG 2018

This document contains privileged and confidential information belonging to the Sponsor. Except as otherwise agreed to in writing, by accepting or reviewing this document, you agree to hold the information contained herein in confidence and not copy or disclose it to others (except where required by applicable law) or use it for unauthorized purposes. In the event information is disclosed to others, as required by applicable law, they must be informed that the information is privileged and confidential and may not be further disclosed by them. These restrictions on disclosures will apply equally to all future information supplied to you that is indicated as privileged and confidential. You agree that no unpublished information contained herein will be published or disclosed without the prior written approval of Sponsor. In the event of any actual or suspected breach of this obligation, Sponsor must be promptly notified.

TABLE OF CONTENTS

1 INTRODUCTION..... 4

1.1 STUDY OBJECTIVES 5

1.2 STUDY DESIGN 5

2 INTERIM ANALYSES 7

3 ANALYSIS SETS..... 8

3.1 FULL ANALYSIS SET 8

3.2 ‘PER PROTOCOL’ ANALYSIS SET 8

3.3 SAFETY ANALYSIS SET 8

3.4 OTHER ANALYSIS SETS..... 8

4 ENDPOINTS AND COVARIATES..... 8

Reflected Change in Nasal Congestion Symptoms 9

Reflected Congestion Severity..... 10

Wisconsin Upper Respiratory Symptom Survey 21 10

EFFICACY ENDPOINT(S)..... 11

4.1.1 Primary endpoint 11

4.1.2 Secondary endpoints 11

4.1.3 Exploratory endpoints..... 11

4.2 SAFETY ENDPOINTS 12

4.3 OTHER ENDPOINTS 12

4.4 COVARIATES 12

5 HANDLING OF MISSING VALUES 13

Deviations in Day 1 post-baseline assessment times..... 13

6 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES..... 13

6.1 STATISTICAL HYPOTHESES 13

6.2 STATISTICAL DECISION RULES 14

6.3 STATISTICAL METHODS..... 14

6.4 STATISTICAL ANALYSES 15

6.4.1 Demographic and Baseline Characteristics 15

6.4.2 Analysis of primary and secondary efficacy endpoints..... 16

6.4.3 Analysis of exploratory efficacy endpoints 16

6.4.4 Safety Analyses..... 16

APPENDIX 1: SUMMARY TABLES 18

APPENDIX 2: SUBJECT DATA LISTINGS 19

APPENDIX 3: FIGURES..... 20

APPENDIX 4: INTERIM STATISTICAL ANALYSIS PLAN..... 21

1 INTRODUCTION

The background and rationale of the current study as stated in the protocol:

Phenylephrine HCl is used as a nasal decongestant in over-the-counter adult and pediatric cough and cold medicines. It is indicated for use for the temporary relief of nasal and sinus congestion and sinus pressure due to the common cold, hay fever, or other upper respiratory allergies. Phenylephrine HCl is commercially available in many countries worldwide as [REDACTED] single ingredient and combination products for oral administration. The dosing regimen for adults and children 12 years of age and older is one dose of phenylephrine HCl 10 mg every four hours over a 24-hour period or 12 mg up to four times daily.

Phenylephrine is rapidly absorbed between 15 and 60 minutes following oral administration and undergoes high first-pass metabolism in the intestinal wall. The bioavailability of phenylephrine after an oral dose is approximately 38%. When phenylephrine HCl is administered after a high-fat meal, the absorption rate is decreased, lowering maximum plasma concentrations about 50%. However, total extent of absorption is not affected by food. Both maximum and total drug exposures increase more than proportionally with increasing single doses from 5 to 30 mg phenylephrine HCl, although the disproportional increase is relatively modest. The terminal elimination $t_{1/2}$ of phenylephrine is short, approximately 1 to 2 hours.

Phenylephrine HCl is available in a variety of [REDACTED] single ingredient and combination products. [REDACTED]

1.1 STUDY OBJECTIVES

As stated in the protocol:

'To investigate the potential for efficacy of [REDACTED] to phenylephrine HCl [REDACTED] and to compare with phenylephrine HCl [REDACTED] 12 mg capsules four times a day and placebo, administered orally, in subjects with nasal congestion due to the common cold.'

1.2 STUDY DESIGN

As extracted from the protocol:

'This will be a randomized, double-blind, placebo controlled, parallel-group Phase 2 study to evaluate the efficacy of [REDACTED] to phenylephrine HCl [REDACTED] and [REDACTED] oral capsule of phenylephrine HCl 12 mg taken four times daily for relief of nasal congestion in subjects with naturally occurring cold symptoms. The trial will have a double dummy design [REDACTED] Subjects randomized to the phenylephrine HCl [REDACTED] [REDACTED] Subjects randomized to the phenylephrine HCl [REDACTED] group will take [REDACTED] four doses of active product. The morning dose will be administered between 7:00AM and noon at the site, the next dose will be administered 4 hours (+10 minutes) later at the site and the last two doses given to the subjects to take 4 hours and 8 hours later. Approximately 450 subjects (150 per treatment group), ages 18 years and older, will be enrolled.

Potential subjects will be identified through advertising and unsolicited presentation at the study site for complaint of common cold symptoms. Eligible subjects will have an onset of cold symptoms (per subject self-report) within approximately 72 hours (3 days) before study entry, consisting of cold symptoms per the following:

- *At Visit 1, at least moderate severity (score ≥ 5 , on a scale from 0 = none to 7 = severe) for stuffy/congested nose, and*
- *At Visit 1, at least mild (score ≥ 3 , on a scale from 0 = none to 7 = severe) for sinus pressure/tenderness, and*
- *Within the past 72 hours, two or more of the following symptoms: runny nose, sore or scratchy throat, sneezing, headache, malaise, or cough*

Subjects who qualify per inclusion and exclusion criteria during onsite screening will be randomly assigned via IWRS to one of three blinded treatments: [REDACTED]

[REDACTED] phenylephrine HCl 30 mg, [REDACTED] phenylephrine HCl 12 mg, or placebo. Treatment will be for a twelve-hour period (Day 1).

[REDACTED]

One third of eligible subjects will be assigned to the pharmacokinetic cohort. They will have 3-mL blood samples collected by direct venipuncture. Two samples will be collected; one at approximately 1 (± 10 minutes) and another at 4 hrs (± 20 minutes) after the initial dose. These subjects will have provided written consent to the collection of blood samples. After the pharmacokinetic cohort has been completed, subsequent enrolling subjects will not be asked to consent to the collection of blood samples.

Subjects will be instructed to complete the final assessments in their diary the next morning on Day 2, 24 hours after the first dose was administered. The 24-hour assessments can also be completed during the follow-up visit if it falls within the window. Subjects will return to the study site for a Follow-up Visit on Day 2 (+1 day) for reconciliation of drug and reporting of any adverse events or concomitant medications. Vital signs will be obtained and the subject will be discharged from the study.

No other cough, cold, allergy, or analgesic/antipyretic medicines (nonprescription or prescription) or herbal/dietary supplements will be permitted during the study. Adverse events will be collected during the study to evaluate safety.'

'Subjects who qualify for study entry per inclusion and exclusion criteria during onsite screening will be randomly assigned to one of three blinded treatments: [REDACTED]

[REDACTED]

Study staff will administer the first dose between approximately 07:00 and 12:00 hours (0 hour) on Day 1 and the second dose 4 hours later. Subjects will be discharged from the site after the second dose and will be instructed to continue taking the IP as directed for the rest of the day.'

The randomization schedule will be generated by the Sponsor. Subjects will be randomized via an IWRS to receive one of three treatments [REDACTED] or placebo, 1:1:1 ratio).

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

2 INTERIM ANALYSES

No interim analysis was originally planned for this study. However, planned enrollment was not achieved during the 2017-2018 cold season as previously planned, which would necessitate continued enrollment into subsequent cold season(s) to complete the study's planned enrollment. In order to avoid exposing patients to unnecessary clinical research, an interim futility analysis has been performed to assess the likelihood of a study result that would lead to an acceptable Phase III sample size estimate.

Only interim data on the primary study endpoint was analyzed for the interim statistical analysis and the analysis was performed by an independent third party statistician and programmer. They are, up to the lock of the completed data base, the only people with access to both the subject data and the treatment assignments.

Detailed methodology for the interim statistical analysis was documented in an Interim Statistical Analysis Plan (see Appendix 4) finalized and approved prior to release of randomization codes to the independent third party.

Study data for the interim analysis is stored in a version-controlled environment, and will not include treatment assignments. Treatment/randomization codes were provided by the JNJ randomization administrator directly to the independent third party statistician and independent programmer who would be performing the analysis. Treatment/randomization codes are stored in a secure area only accessible to the independent statistician and independent programmer, according to the third party SOPs.

Upon completion of the analysis, the independent third party statistician communicated results directly and restricted (via secure email exchange) to key stakeholders within the sponsor organization, that is, the Vice President of Over-the-Counter Research & Development and the Vice President of Global Medical Affairs. These individuals could provide the results to the Chief Technology Officer of Consumer and the President of Global Franchise Organization, Worldwide Over-the-Counter Marketing. Results did not include identification of subject treatment assignments.

3 ANALYSIS SETS

3.1 FULL ANALYSIS SET

For the efficacy analyses, the Full Analysis Set (FAS) will include all randomized subjects who provide a valid baseline assessment of nasal congestion severity.

3.2 'PER PROTOCOL' ANALYSIS SET

No Per Protocol Analysis will be performed.

3.3 SAFETY ANALYSIS SET

Safety Analysis Set: The safety analyses will be based on all randomized subjects who take at least one dose of treatment.

3.4 OTHER ANALYSIS SETS

None.

4 ENDPOINTS AND COVARIATES

Efficacy Assessments (As stated in the protocol)

The decongestant effect of phenylephrine [REDACTED] will be evaluated with the following subjective assessment questions and Verbal Rating Scales (VRS) at the designated time points using the Diary:

- *Assessment of Nasal Congestion (stuffy/congested nose), instantaneous, on an 8-point scale, at baseline, at 2, 4, 6, 8, 10 and 12 hours after the first dose on Day 1, and on the morning of Day 2 (24 hours)*
- *Assessment of Sinus Pressure/Tenderness, instantaneous, on an 8-point scale, at baseline, at 2, 4, 6, 8, 10 and 12 hours after the first dose on Day 1, and on the morning of Day 2 (24 hours)*
- *Assessment of Head Congestion, instantaneous, on an 8-point scale, at baseline, at 2, 4, 6, 8, 10 and 12 hours after the first dose on Day 1, and on the morning of Day 2 (24 hours)*
- *Assessment of Reflected Impact of Nasal / Sinus Congestion on Clear Thinking, on a 5-point scale, at baseline and at 12 hours after the first dose on Day 1.*
- *Assessments of Reflected Change in Nasal Congestion Symptoms reflective, on a 7-point scale, at 12 hours after the morning dose on Day 1.*

- *Assessments of Reflected Nasal Functioning (clear breathing), reflective, on an 8-point scale, at 12 hours after the morning dose on Day 1.*
- *Assessment of Reflected Congestion Severity, reflective, on a 5-point scale, at baseline and at 12 hours after the morning dose on Day 1.*
- *Assessment of cold severity using the Wisconsin Upper Respiratory Symptom Survey (WURSS-21) at baseline and on the morning of Day 2 (24 hours).'*

Rating scales used:

Nasal Congestion, Sinus Pressure and Head Congestion:

Rate the severity of each symptom at this point in time	None	Very Mild		Mild		Moderate		Severe
	0	1	2	3	4	5	6	7
Stuffy / congested nose	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reflected Impact of Nasal / Sinus Congestion on Clear Thinking

To what extent has your nasal/sinus congestion affected your ability to think clearly over the past 12 hours?	No effect on clear thinking	Very mild	Mild	Moderate	Severe
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reflected Change in Nasal Congestion Symptoms

How are your nasal congestion symptoms now, compared to just before you were given the first dose of study medication at the start of the study?	Much worse	Worse	A little worse	No change	A little better	Better	Much better
	0	1	2	3	4	5	6
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reflected Nasal Functioning

When breathing, how clear has your nose felt over the whole day since the first dose of study medication?	Clear	Mostly clear		Partly clear		A little clear		Not clear at all
	0	1	2	3	4	5	6	7
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reflected Congestion Severity

Rate the average severity of your stuffy nose over the past 12 hours?	No stuffy nose	Very mild	Mild	Moderate	Severe
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Wisconsin Upper Respiratory Symptom Survey 21

Wisconsin Upper Respiratory Symptom Survey – 21 --- Daily Symptom Report

Day: _____ Date: _____ Time: _____ ID: _____

Please fill in one circle for each of the following items:

	Not sick 0	Very mildly 1	Mildly 2	Moderately 3	Moderately 4	Severely 5	Severely 6	Severely 7
How sick do you feel today?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please rate the average severity of your cold symptoms over the last 24 hours for each symptom:

	Do not have this symptom 0	Very mild 1	Mildly 2	Mildly 3	Moderate 4	Moderate 5	Severe 6	Severe 7
Runny nose	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Plugged nose	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sneezing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sore throat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scratchy throat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cough	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hoarseness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Head congestion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Chest congestion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Feeling tired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Over the last 24 hours, how much has your cold interfered with your ability to:

	Not at all 0	Very mildly 1	Mildly 2	Mildly 3	Moderately 4	Moderately 5	Severely 6	Severely 7
Think clearly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sleep well	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Breathe easily	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Walk, climb stairs, exercise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Accomplish daily activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Work outside the home	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Work inside the home	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interact with others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Live your personal life	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Compared to yesterday, I feel that my cold is...

Very much better	Somewhat better	A little better	The same	A little worse	Somewhat worse	Very much worse
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

WURSS -21® (Wisconsin Upper Respiratory Symptom Survey) 2004
 Created by Bruce Barrett MD PhD et al., UW Department of Family Medicine, 777 S. Mills St. Madison, WI 53715, USA

EFFICACY ENDPOINT(S)

4.1.1 Primary endpoint

The primary efficacy endpoint is the change from baseline (score improvement) in the nasal congestion severity score (NCSS) averaged over assessments at 2, 4, 6, 8, 10, and 12 hours on Day 1. The baseline nasal congestion score will be measured within approximately 30 minutes before the first dose.

4.1.2 Secondary endpoints

- Change from baseline (score improvement) in the NCSS averaged over assessments at 8, 10, and 12 hours.
- Change from baseline (score improvement) in the NCSS at 2, 4, 6, 8, 10, and 12 hours and in the morning of Day 2 (24 hours), respectively.
- Change from baseline (score improvement) in Sinus Pressure/Tenderness Scores averaged over assessments at 2, 4, 6, 8, 10, and 12 hours.
- Change from baseline (score improvement) in Sinus Pressure/Tenderness Scores at 2, 4, 6, 8, 10, and 12 hours and in the morning of Day 2 (24 hours), respectively.

4.1.3 Exploratory endpoints

- Change from baseline (score improvement) in Head Congestion Scores averaged over assessments at 2, 4, 6, 8, 10, and 12 hours.
- Change from baseline (score improvement) in Head Congestion Scores at 2, 4, 6, 8, 10, and 12 hours and in the morning of Day 2 (24 hours), respectively.
- Change from baseline (score improvement) in Reflected Impact of Nasal / Sinus Congestion on Clear Thinking (reflected for the previous 12 hours) at 12 hour assessment
- Reflected Change in Nasal Congestion Symptoms (reflected over 12 hours after the first dose).
- Reflected Nasal Functioning (reflected over 12 hours after the first dose).
- Change from baseline (score improvement) in Reflected Congestion Severity (reflected for the previous 12 hours) at 12 hour assessment.
- Assessment of cold severity using Wisconsin Upper Respiratory Symptom Survey-21 (reflected over 24 hours) at baseline and in the morning of Day 2 (24 hours).

4.2 SAFETY ENDPOINTS

- Occurrence of self-reported and observed adverse events.
- Occurrence of serious adverse events.
- Occurrence of adverse events leading to subject discontinuation from the study.
- The number and percentage of subjects experiencing treatment-emergent adverse events.
- The number and percentage of subjects experiencing treatment related adverse events, i.e. adverse events with a possible, probable, or very likely relation to the investigational product.
- Heart rate (pulse) and systolic and diastolic blood pressure data at baseline (pre-dose) and at 2 and 4 hours post-dose, immediately following completion of subject assessments, as well as at end of study (Day 2).
- Respiratory rate and oral body temperature measured at baseline (pre-dose) and end of study (Day 2).

4.3 OTHER ENDPOINTS

Data collected in this study will also be used to evaluate the psychometric properties of the primary and secondary outcome assessments in this context of use and to perform analyses to estimate the level of change in scores that can be considered meaningful and important. These analyses will be described in a specific and separate psychometric analysis plan which will be finalized prior to database lock and the analysis will be reported separately. The details of the psychometric data analyses are not described in this document.

No pharmacokinetic analysis will be performed for the current study because the clinical program associated with the study has been discontinued and therefore no additional pharmacokinetic data is required to support future clinical study designs.

4.4 COVARIATES

ANOVA models for analyses of change from baseline in the NCSS scores will be adjusted for the baseline NCSS score categorized as either 'equal to 5' or 'greater than 5'. (A blind data review found only 14 (7.25%) enrolled subjects with a baseline NCSS score equal to 7, hence the merger between the two highest score categories, '6' and '7'.)

ANCOVA models for analyses of change from baseline in Sinus Pressure/Tenderness Scores and Head Congestion scores, respectively, will be adjusted for the corresponding baseline score.

5 HANDLING OF MISSING VALUES

In the analyses of change from baseline in the NCSS, Sinus Pressure/Tenderness Scores, and Head Congestion scores, respectively, subjects who dropout, will have their last reported corresponding score before the dropout imputed for the remaining assessment time points. Any intermittent missing assessments will be imputed using linear interpolation between adjacent assessments.

The analyses of score changes from baseline in Reflected Impact of Nasal / Sinus Congestion on Clear Thinking, Reflected Change in Nasal Congestion Symptoms scores, Reflected Nasal Functioning scores, and score changes from baseline in Reflected Congestion Severity will be restricted to complete cases, that is, subjects with non-missing data for the corresponding endpoint.

Deviations in Day 1 post-baseline assessment times

For time deviations in Day 1 post-baseline assessment time points exceeding 10 minutes in either direction relative to target, the corresponding assessment values will be adjusted using linear interpolation and the interpolated value at the target assessment time point will be used in the statistical analyses. Any 12 hour assessments performed more than 10 minutes before target, will be adjusted by linear extrapolation to 12 hours using the 10 and 12 hour assessments.

6 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

6.1 STATISTICAL HYPOTHESES

This is a Phase II, Proof-Of-Concept (POC) study and the inferential statistical procedures are focused on point and interval estimation of the mean treatment differences in efficacy endpoints based on changes from baseline in the NCSS, Sinus Pressure/Tenderness Scores, and Head Congestion scores, respectively. However, to further aid interpretation of results, p-values from 2-sided tests of the null hypotheses of no treatment mean differences will be provided. Specifically, for each of the endpoints derived from these scores and each pair of treatments the following null hypothesis will be tested,

$$H_0: \mu_1 = \mu_2$$

versus the alternative hypothesis,

$$H_a: \mu_1 \neq \mu_2$$

where μ_1 and μ_2 are the underlying mean endpoint values of the two compared treatments.

The analyses of score changes from baseline in Reflected Impact of Nasal / Sinus Congestion on Clear Thinking, Reflected Change in Nasal Congestion Symptoms scores, Reflected Nasal Functioning scores, and score changes from baseline in Reflected Congestion Severity, respectively, will for each separate endpoint and pair of treatments test the null hypothesis of no treatment difference in the endpoint distributions. That is,

H_0 : The endpoint distributions are the same for both compared treatments

versus the alternative hypothesis,

H_1 : The endpoint distributions differ for the two compared treatments

6.2 STATISTICAL DECISION RULES

This is a Phase II, POC study and the inferential statistical procedures are focused on point and interval estimation of the mean treatment differences in efficacy endpoints based on changes from baseline in the NCSS, Sinus Pressure/Tenderness Scores, and Head Congestion scores, respectively. All interval estimates will be two-sided and have nominal confidence level 95%.

6.3 STATISTICAL METHODS

In the analyses of change from baseline in the NCSS, Sinus Pressure/Tenderness Scores, and Head Congestion scores, respectively, subjects who dropout, will have their last reported corresponding score before the dropout imputed for the remaining assessment time points. The analyses will be based on all randomized subjects who provides at least a baseline assessment and receives study treatment.

The primary endpoint, which is the overall Day 1 change from baseline in the Nasal Congestion Score (score for stuffy/congested nose averaged over assessments at 2, 4, 6,

8, 10, and 12 hours on Day 1), will be analyzed in an ANOVA model with treatment group, and baseline nasal congestion score ($5, \geq 6$) as factors.

Analysis of all secondary endpoints involving instantaneous nasal congestion scores will be performed using an ANOVA model with treatment group, and baseline nasal congestion score ($5, \geq 6$) as factors.

Analysis of secondary endpoints involving Sinus Pressure/Tenderness Scores will be performed using an ANCOVA model with treatment group as factor, and baseline Sinus Pressure/Tenderness Score as a covariate.

Analysis of endpoints involving Head Congestion Scores will be performed using an ANCOVA model with treatment group as factor, and baseline Head Congestion Score as a covariate.

Pairwise treatment differences in the distributions of Reflected Change in Nasal Congestion Symptoms and Reflected Nasal Functioning scores, both reflected over 12 hours after the first dose, will be analyzed using the Mann-Whitney test.

Pairwise treatment differences in the distributions of Change from baseline in Reflected Impact of Nasal / Sinus Congestion on Clear Thinking and Change from baseline in Reflected Congestion Severity (reflected for the previous 12 hours), will be analyzed using the Mann-Whitney test.

All efficacy endpoints will be described by summary statistics. This includes summaries of the first and last questions ('How sick do you feel today' and 'Compared to yesterday, my cold is ...') in the assessments of cold severity using the Wisconsin Upper Respiratory Symptom Survey (WURSS-21) at baseline and in the morning of Day 2 (24 hours), as well as descriptive statistics for a summary score derived from the remaining 19 questionnaire items.

All presented confidence intervals will be two-sided and have confidence level 95%. To further guide data interpretation, p-values from 2-sided tests of the null hypotheses specified in Section 6.1 will be presented.

6.4 STATISTICAL ANALYSES

The Department of Quantitative Science at McNeil AB will be responsible for the statistical data analyses. SAS version 9.4 will be used for the statistical analyses.

6.4.1 Demographic and Baseline Characteristics

Descriptive statistics (e.g. mean, standard deviation, median, minimum and maximum for continuous variables; frequency and percentage for categorical variables) will be presented for demographic and baseline characteristics.

6.4.2 Analysis of primary and secondary efficacy endpoints

Descriptive statistics of primary and secondary efficacy endpoints will be presented based on all subjects included in the analyses, see 6.3.

The comparative statistical analyses will be performed as described in Section 6.3 and will be based on imputations of any missing scores or out-of-window data, see Section 5.

This is an exploratory Phase 2 POC study. Additional post-hoc analyses may be performed.

6.4.3 Analysis of exploratory efficacy endpoints

Descriptive statistics of exploratory endpoints will be presented based on all subjects in the Full Analysis Set with non-missing endpoint values (no imputation performed). The comparative statistical analyses will be performed as described in Section 6.3. The only exception concerns endpoints based on Head Congestion scores which will be summarized and analyzed according to the description in 6.4.2.

This is an exploratory Phase 2 POC study. Additional post-hoc analyses may be performed.

6.4.4 Safety Analyses

Safety analyses will be based on the Safety Analysis Set.

All adverse events reported during the Adverse Event reporting period will be listed by subject ID. Any serious adverse events will be listed separately.

The number and percentage of subjects experiencing adverse events will be tabulated by treatment, system organ class, and preferred term. Separate tables will be provided for treatment-emergent and treatment-related adverse events, the latter being restricted to adverse events having a possible, probable, or very likely relation to the investigational product. Any adverse events occurring before the first treatment administration will be listed only. (Any adverse events with onset date coinciding with the date of the first treatment administration but missing time of onset will be considered as treatment-emergent in tables and listings of adverse events.)

In addition, the number and percentage of subjects experiencing adverse events will also be tabulated by treatment, system organ class, preferred term, and worst recorded severity. Separate tables will be provided for treatment-emergent and treatment-related adverse events.

Any subjects who discontinued the study due to a treatment-emergent adverse event(s) will also be separately listed by treatment including descriptions of the adverse events leading to withdrawal.

The latest version of the Medical Dictionary for Regulatory Activities (MedDRA) will be used as Adverse Event classification system.

Heart rate (pulse) and systolic and diastolic blood pressure data summaries will be presented by treatment and assessment time point for assessments at baseline (pre-dose) and at 2 and 4 hours post-dose, immediately following completion of subject assessments.

Respiratory rate, oral body temperature, heart rate (pulse) and systolic and diastolic blood pressure data measured at baseline (pre-dose) and end of study (Day 2) will be presented by treatment and visit.

Previous and Concomitant medications will be listed only.

APPENDIX 1: SUMMARY TABLES

- 14.1.1 Disposition of Subjects. All randomized subjects.
- 14.1.2 Demography and Baseline Characteristics. All randomized subjects.
- 14.1.3 Vital signs at Screening. All randomized subjects.
- 14.2.1.1 Change from baseline in the nasal congestion severity score (NCSS) by assessment time point. Subjects in Full Analysis Set.
- 14.2.1.2 Mean Change from baseline in the nasal congestion severity score (NCSS). Subjects in Full Analysis Set.
- 14.2.2.1 Change from baseline in Sinus Pressure/Tenderness Scores by assessment time point. Subjects in Full Analysis Set.
- 14.2.2.2 Mean Change from baseline in Sinus Pressure/Tenderness Scores. Subjects in Full Analysis Set.
- 14.2.3.1 Change from baseline in Head Congestion Scores by assessment time point. Subjects in Full Analysis Set.
- 14.2.3.2 Primary analysis of Mean Change from baseline in Head Congestion Scores. Subjects in Full Analysis Set.
- 14.2.4 Change from baseline in Reflected Impact of Nasal / Sinus Congestion on Clear Thinking (reflected for the previous 12 hours). Subjects in Full Analysis Set.
- 14.2.5 Reflected Change in Nasal Congestion Symptoms (reflected over 12 hours after the first dose). Subjects in Full Analysis Set.
- 14.2.6 Reflected Nasal Functioning (reflected over 12 hours after the first dose). Subjects in Full Analysis Set.
- 14.2.7 Change from baseline (score improvement) in Reflected Congestion Severity (reflected for the previous 12 hours). Subjects in Full Analysis Set.
- 14.2.8 Assessment of cold severity using Wisconsin Upper Respiratory Symptom Survey-21 (reflected over 24 hours). Subjects in Full Analysis Set.
- 14.2.9 Compliance with investigational product. All randomized subjects.
- 14.3.1.1 Subjects with Treatment Emergent Adverse Events By System Organ Class and Preferred Term. Subjects in Safety Analysis Set.

- 14.3.1.2 Subjects with Commonly Reported ($\geq 5\%$) Treatment Emergent Adverse Events By System Organ Class and Preferred Term. Subjects in Safety Analysis Set.
- 14.3.1.3 Subjects with Treatment Emergent Adverse Events By System Organ Class, Preferred Term and Severity. Subjects in Safety Analysis Set.
- 14.3.1.4 Subjects with Treatment Related Adverse Events By System Organ Class and Preferred Term. Subjects in Safety Analysis Set.
- 14.3.1.5 Subjects with Treatment Related Adverse Events By System Organ Class, Preferred Term and Severity. Subjects in Safety Analysis Set.
- 14.3.1.6 Vital Signs at Baseline and End-of-Study. Subjects in Safety Analysis Set.
- 14.3.1.7 Vital Signs Pre and Post Treatment. Subjects in Safety Analysis Set.
- 14.3.2 Listing of Deaths, Other Serious and Significant Adverse Events. All randomized subjects.

APPENDIX 2: SUBJECT DATA LISTINGS

- Randomization code listing. (16.1.7)
- Study disposition including status at termination. (16.2.1)
- Protocol deviations. (16.2.2)
- Subjects excluded from the efficacy analysis. (16.2.3)
- Demographic and baseline characteristics. (16.2.4.1)
- Vital signs at Screening. (16.2.4.2)
- Subjects with Medical History. (16.2.4.3)
- Compliance. (16.2.5)
- Nasal congestion severity scores (NCSS) by assessment time points. (16.2.6.1.1)
- Nasal congestion severity scores (NCSS). Mean change from baseline. (16.2.6.1.2)
- Sinus Pressure/Tenderness Scores by assessment time points. (16.2.6.2.1)
- Sinus Pressure/Tenderness Scores. Mean change from baseline. (16.2.6.2.2)

- Head Congestion Scores by assessment time points. (16.2.6.3.1)
- Head Congestion Scores. Mean change from baseline. (16.2.6.3.2)
- Reflected Impact of Nasal / Sinus Congestion on Clear Thinking, and Reflected Congestion Severity. (16.2.6.4)
- Reflected Change in Nasal Congestion Symptoms, and Reflected Nasal Functioning. (16.2.6.5)
- Wisconsin Upper Respiratory Symptom Survey-21. (16.2.6.5)
- Adverse events. (16.2.7.1)
- Vital Signs. (16.2.7.2)
- Previous and Concomitant medication. (16.2.9)

APPENDIX 3: FIGURES

1. Nasal Congestion Severity Scores Day 1 (Mean \pm 2SE).
2. Sinus Pressure/Tenderness Scores Day 1 (Mean \pm 2SE).
3. Head Congestion Scores Day 1 (Mean \pm 2SE).

APPENDIX 4: INTERIM STATISTICAL ANALYSIS PLAN

Interim Statistical Analysis Plan

***RANDOMIZED, DOUBLE-BLIND, PLACEBO-
CONTROLLED, STUDY OF THE EFFICACY OF
PHENYLEPHRINE HCL [REDACTED]
MG AND PHENYLEPHRINE HCL [REDACTED]
[REDACTED] 12 MG CAPSULES IN SUBJECTS WITH
NASAL CONGESTION DUE TO THE COMMON
COLD***

<i>Investigational Product Name:</i>	<i>Phenylephrine HCl</i> [REDACTED]
<i>Protocol Number:</i>	<i>CO-170302131230-URCT</i>
<i>IND / EudraCT number:</i>	<i>IND: 121517</i>
<i>Phase:</i>	<i>2</i>
<i>Version: Date</i>	<i>FINAL: 26 JUN 2018</i>

1 RATIONALE

The background and rationale of the interim analysis as stated in the amended protocol:

'Planned enrollment was not achieved during the 2017-2018 cold season as previously planned, necessitating continued enrollment into subsequent cold season(s) to complete the study's planned enrollment. In order to avoid exposing patients to unnecessary clinical research, an interim futility analysis will occur to assess the likelihood of a study result that would lead to an acceptable Phase III sample size estimate.

The potential acceptability of the future Phase III study will be directly proportional to the sample size/sponsor resourcing required for the study. The sponsor has not defined an acceptable sample size/resourcing limit. Therefore, the interim analysis will include three sample size thresholds.'

2 INTERIM ANALYSIS COMMUNICATIONS AND LOGISTICS

As stated in the amended protocol:

'The statistical analysis of interim study data will be performed by an independent third party statistician and programmer. They will be the only people who will have access to both the subject data and the treatment assignments.

Detailed methodology for the interim statistical analysis will be documented in an Interim Statistical Analysis Plan finalized and approved prior to release of randomization codes to the independent third party.

Study data for the interim analysis will be stored in a version-controlled environment, and will not include treatment assignments. Treatment/randomization codes will be provided by the JNJ randomization administrator directly to the independent third party statistician and independent programmer who will be performing the analysis. Treatment/randomization codes will be stored in a secure area only accessible to the independent statistician and independent programmer, according to the third party SOPs.

Upon completion of the analysis, the independent third party statistician will communicate results directly and restricted (via secure email exchange) to key stakeholders within the sponsor organization, that is, the Vice President of Over-the-Counter Research & Development and the Vice President of Global Medical Affairs. These individuals may provide the results to the Chief Technology Officer of Consumer and the President of Global Franchise Organization, Worldwide Over-the-Counter Marketing. Results will not include identification of subject treatment assignments.'

3 STUDY OBJECTIVES AND DESIGN

3.1 STUDY OBJECTIVES

As stated in the protocol:

'To investigate the potential for efficacy of [REDACTED] to phenylephrine HCl 30 mg [REDACTED] and to compare with phenylephrine HCl [REDACTED] 12 mg capsules four times a day and placebo, administered orally, in subjects with nasal congestion due to the common cold.'

3.2 STUDY DESIGN

As extracted from the protocol:

'This will be a randomized, double-blind, placebo controlled, parallel-group Phase 2 study to evaluate the efficacy of [REDACTED] to phenylephrine HCl [REDACTED] and [REDACTED] oral capsule of phenylephrine HCl 12 mg taken four times daily for relief of nasal congestion in subjects with naturally occurring cold symptoms. [REDACTED]

Potential subjects will be identified through advertising and unsolicited presentation at the study site for complaint of common cold symptoms. Eligible subjects will have an onset of cold symptoms (per subject self-report) within approximately 60 hours (2.5 days) before study entry, consisting of at least four cold symptoms per the following:

- At Visit 1, at least moderate severity (score ≥ 5 , on a scale from 0 = none to 7 = severe) for stuffy/congested nose;*
- At Visit 1, at least mild (score ≥ 3 , on a scale from 0 = none to 7 = severe) for sinus pressure/tenderness;*
- Within the past 60 hours, at least two of the following symptoms: runny nose, sore or scratchy throat, sneezing, headache, malaise, or cough*

Subjects who qualify per inclusion and exclusion criteria during onsite screening will be randomly assigned via IWRS to one of three blinded treatments: [REDACTED] phenylephrine HCl 30 mg, [REDACTED] phenylephrine HCl 12 mg, or placebo. Treatment will be for a twelve-hour period (Day 1).

On Day 1, subjects will complete subjective assessments of symptoms at baseline, within [REDACTED]

One third of eligible subjects will be assigned to the pharmacokinetic cohort. They will have 3-mL blood samples collected by direct venipuncture. Two samples will be collected; one at approximately 1 (± 10 minutes) and another at 4 hrs (± 20 minutes) after the initial dose. These subjects will have provided written consent to the collection of blood samples. After the pharmacokinetic cohort has been completed, subsequent enrolling subjects will not be asked to consent to the collection of blood samples.

Subjects will be instructed to complete the final assessments in their diary the next morning on Day 2, 24 hours after the first dose was administered. Subjects will return to the study site for a Follow-up Visit on Day 2 (+1 day) for reconciliation of drug and reporting of any adverse events or concomitant medications. Vital signs will be obtained and the subject will be discharged from the study.

No other cough, cold, allergy, or analgesic/antipyretic medicines (nonprescription or prescription) or herbal/dietary supplements will be permitted during the study. Adverse events will be collected during the study to evaluate safety.'

'Subjects who qualify for study entry per inclusion and exclusion criteria during onsite screening will be randomly assigned to one of three blinded treatments: [REDACTED]

[REDACTED] or placebo 4 times a day, for one day.

Study staff will administer the first dose between approximately 08:00 and 11:00 hours (0 hour) on Day 1 and the second dose 4 hours later. Subjects will be discharged from the site after the second dose and will be instructed to continue taking the IP as directed for the rest of the day'

The randomization schedule was generated by the Sponsor. Subjects were randomized via an IWRS to receive one of three treatments [REDACTED] or placebo, 1:1:1 ratio).

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

4 SCOPE OF INTERIM ANALYSIS

The interim statistical analysis will be an analysis of futility based exclusively on interim data on the study primary endpoint, that is, the change from baseline (score improvement) in the nasal congestion severity score (NCSS) averaged over assessments at 2, 4, 6, 8, 10, and 12 hours on Day 1.

5 ANALYSIS SETS

5.1 FULL ANALYSIS SET

For the interim statistical analysis of the primary study endpoint, the Full Analysis Set (FAS) will include all randomized subjects who provided a valid baseline assessment of nasal congestion severity.

5.2 'PER PROTOCOL' ANALYSIS SET

No Per Protocol Interim Analysis will be performed.

5.3 SAFETY ANALYSIS SET

No Safety Interim Analysis will be performed.

5.4 OTHER ANALYSIS SETS

Not applicable.

6 ENDPOINTS AND COVARIATES

6.1 EFFICACY ENDPOINT

The only endpoint that will be addressed in the interim analysis is the study primary efficacy endpoint, that is, the change from baseline (score improvement) in the nasal congestion severity score (NCSS) averaged over assessments at 2, 4, 6, 8, 10, and 12 hours on Day 1. Instantaneous assessments of nasal congestion (stuffy/congested nose), were made on an 8-point scale, at baseline, at 2, 4, 6, 8, 10 and 12 hours after the first dose on Day 1:

Rate the severity of each symptom at this point in time	None	Very Mild		Mild		Moderate		Severe
	0	1	2	3	4	5	6	7
Stuffy / congested nose	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6.2 SAFETY ENDPOINTS

Not applicable.

6.3 OTHER ENDPOINTS

Not applicable.

6.4 COVARIATES

None.

7 HANDLING OF MISSING VALUES AND ASSESSMENT TIME DEVIATIONS

In the interim futility analysis of the primary study endpoint, subjects who were withdrawn any time after the baseline NCSS assessment, will have their last reported NCSS before the dropout imputed for the remaining assessment time points during Day 1. For example, any subjects without post-baseline assessments will have their baseline values carried forward. Any intermittent missing assessments will be imputed using linear interpolation between adjacent assessments.

Post-baseline NCSS assessments with time deviations not exceeding 10 minutes in either direction relative to target, will be treated in endpoint derivations as having occurred on target with no adjustments of the corresponding assessment values. For time deviations in post-baseline assessment time points exceeding 10 minutes in either direction relative to target, the corresponding assessment values at the target time points will be estimated

using linear interpolation using adjacent assessment values at target time points. If the actual 12 hour assessment is made more than 10 minutes before the target time point, the assessment value at 12 hours will be estimated using extrapolation based on the last two assessments preceding 12 hours. The interpolated and extrapolated assessment values will be used in the statistical interim analysis.

8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 STATISTICAL METHODS

The interim futility analysis will be based on an adaptation of the methodology described in [1] to the current Phase II study and its objective to serve as a basis for the planning of a confirmatory Phase III non-inferiority study involving the same three treatment groups. The adaptation will be based on the following definition of a successful outcome of a completed Phase II study in the originally targeted 450 subjects (with 150 subjects in each treatment group):

Definition of a Successful Outcome of a completed Phase II POC (SPOC) study:

A completed Phase II study, with 450 randomized subjects enrolled in total from the two study phases, will provide estimates of the population mean differences for the primary study endpoint, [REDACTED] vs. Placebo and [REDACTED] vs. Placebo, respectively, of magnitudes such that a sample size calculation for a Phase III non-inferiority study, based on the data from the completed Phase II study, will suggest a confirmatory Phase III study of total sample size at most N , assuming an equal number of subjects allocated to the three treatment arms, [REDACTED]/Placebo [REDACTED]/Placebo [REDACTED] and Placebo [REDACTED]/Placebo [REDACTED], respectively.

The interim analysis will evaluate futility for three (3) different suggested maximum sample sizes, N , of a confirmatory Phase III non-inferiority study: $N= 1800, 2700, \text{ and } 3600$, respectively.

For each fixed choice of N , two conditional probabilities of a SPOC, defined above, will be calculated given interim data, corresponding to two, somewhat different, approaches to determine the size of a Phase III non-inferiority study:

- a. Based on a pre-set non-inferiority margin derived from an interval estimate of half the population difference in means between [REDACTED] and Placebo (using data from the completed POC study), while assuming the effects of [REDACTED] to be the same and a required statistical power of 90%. This approach assesses the reference product effect and the resulting margin.
- b. Based on a requirement to show, with 90% statistical power, that the population difference in means between [REDACTED] is at least as large as minus half the population difference in means between [REDACTED] Placebo (using data from the completed POC study). This approach assesses the effect of the test product vs. placebo and the reference.

The first preparatory step of the calculation of the conditional probability of a SPOC corresponding to A is to determine a minimal value of the pre-set margin, Δ say, that will correspond to a suggested total Phase III study sample size of at most N subjects under an assumption of a common standard deviation between the three treatment groups and a one-sided significance level of 2.5%. The common standard deviation will be estimated from a pooled variance estimate based on all evaluable interim data. Using normal approximation in the derivation,

$$\Delta = \Delta(N) = \sqrt{6} \times \sigma^* \times (z_{0.025} + z_{0.10}) / \sqrt{N},$$

where σ^* is the pooled standard deviation estimate and z_α is a percentile of the standard normal distribution defined by $1 - \Phi(z_\alpha) = \alpha$.

The second approach (B) does not use a pre-set margin but is instead based on a Phase III non-inferiority evaluation in which one compares the difference in means between [REDACTED] to minus half the difference in means between [REDACTED] and Placebo. In this case the sample size derivation does not assume the effects of [REDACTED] to be the same. With [REDACTED] and μ_{PLACEBO} denoting the population primary endpoint means corresponding to the three treatments, the first preparatory step of the calculation of the corresponding conditional probability of a SPOC is to determine the minimal (population) value of [REDACTED], Δ' say, for which the power to reject the null hypothesis $H_0: [REDACTED] / 2 \leq 0$ is at least 90% with a study in at most N subjects, again under an assumption of a common standard deviation between the three treatment groups and using a one-sided significance level of 2.5%. The common standard deviation will be estimated from a pooled variance estimate based on all evaluable interim data, that is the same estimate as in the first approach will be applied. Using normal approximation in the derivation,

$$\Delta' = \Delta'(N) = 3 \times \sigma^* \times (z_{0.025} + z_{0.10}) / \sqrt{(2N)},$$

where again σ^* is the pooled standard deviation estimate and z_α is a percentile of the standard normal distribution defined by $1 - \Phi(z_\alpha) = \alpha$.

Based on the methodology outlined in [1], for a fixed maximum total sample size N , the calculations of the two conditional probabilities of a successful outcome of a completed POC study, as defined above and corresponding to A and B, are similar. With δ^* denoting the estimate from interim data of half the mean difference between [REDACTED] vs. Placebo, the first conditional probability of a SPOC, corresponding to A, is given by

$$1 - \Phi((z_{0.65} - Z) / \sqrt{f})$$

with $Z = [2 \times (\delta^* - \Delta) \times \sqrt{300}] / [\sigma^* \times \sqrt{(2 + i/j + j/i)}]$ and where i is the number of evaluable subjects in the Placebo arm at interim, j is the number of evaluable subjects in the [REDACTED] arm at interim, $f = 1 - (i+j)/300$, $\Phi(\cdot)$ is the standard normal distribution function, z_α is the percentile defined by $1 - \Phi(z_\alpha) = \alpha$, and σ^* is a pooled standard deviation estimate based on all evaluable interim data. The use of $z_{0.65}$ in the above formula for the conditional probability corresponds to using the lower bound of a one-sided 65% confidence

interval, based on the completed POC study, as an estimate for the margin in a Phase III non-inferiority study sample size calculation according to A.

Similarly for step B, with δ^* denoting the estimate from interim data of $(\bar{x}_{\blacksquare} - \bar{x}_{\blacksquare})/2$, the second conditional probability is given by

$$1 - \Phi\left(\frac{z_{0.50} - Z'}{\sqrt{f'}}\right)$$

where $Z' = [(\delta^* - \Delta) \times \sqrt{450}] / [\sigma^* \times \sqrt{(1.5 + (i+j)/k + (j+k)/4i + (i+k)/4j)}]$ with notation as above extended to: k the number of evaluable subjects in the \blacksquare arm at interim and $f' = 1 - (i+j+k)/450$. The use of $z_{0.50} = 0$ in the above formula corresponds to using the observed point estimate of $(\bar{x}_{\blacksquare} - \bar{x}_{\blacksquare})/2$, based on the completed POC study, in a Phase III non-inferiority study sample size calculation according to B.

Based on the interim data, the estimates σ^* , δ^* and δ'^* will, with obvious notation, be derived according to

$$\sigma^* = \sqrt{[(i-1) \times s^2_{\text{Placebo}} + (j-1) \times s^2_{\blacksquare} + (k-1) \times s^2_{\blacksquare}] / (i+j+k-3)};$$

$$\delta^* = (\bar{x}_{\blacksquare} - \bar{x}_{\text{Placebo}})/2;$$

$$\delta'^* = (\bar{x}_{\blacksquare} - \bar{x}_{\blacksquare}) + (\bar{x}_{\blacksquare} - \bar{x}_{\text{Placebo}})/2 = \bar{x}_{\blacksquare} - (\bar{x}_{\blacksquare} + \bar{x}_{\text{Placebo}})/2.$$

8.2 STATISTICAL DECISION RULES

For each choice of N , a recommendation to stop the study for futility will be given, based on the outcome of the interim statistical analysis, if at least one of the two conditional probabilities of a successful, completed POC study is below 0.2. The only information disseminated to sponsor will be in the form of a 'Yes/No' answer to the question

'Does the interim analysis outcome suggest that the current study should be stopped for futility based on the probability of a maximum tolerable Phase III non-inferiority sample size of, in total, N subjects?'

This interim analysis is a futility analysis. To limit the risk of introducing bias in estimates based on a potential completion of the Phase II study in 450 subjects, the cut-off probability for futility used for evaluation of the conditional probabilities corresponding to the two steps, A and B, have been pre-set to a reasonably low value = 0.2. There will be no other stopping rules applied.

8.3 STATISTICAL ANALYSIS

The Department of Quantitative Science at McNeil AB will be responsible for the statistical data analyses. However, the analysis will be performed by an independent third party statistician and programmer. They will be the only people who will have

access to both the subject data and the treatment assignments. SAS version 9.4 will be used for the statistical analyses.

8.3.1 Futility interim analysis of the study primary endpoint

The interim statistical analysis will be performed as described in Section 8.1. The analysis can, in summary, be described according to the following calculation steps:

$$1. \quad \sigma^* = \sqrt{[(i-1) \times s^2_{Placebo} + (j-1) \times s^2_{\blacksquare} + (k-1) \times s^2_{\blacksquare}] / (i+j+k-3)};$$

For each of $N = 1800, 2700,$ and 3600 then calculate:

$$2.1 \quad \Delta = \Delta(N) = \sqrt{6} \times \sigma^* \times (z_{0.025} + z_{0.10}) / \sqrt{N};$$

$$2.2 \quad \delta^* = (\bar{x}_{\blacksquare} - \bar{x}_{Placebo}) / 2;$$

$$2.3 \quad Z = [2 \times (\delta^* - \Delta) \times \sqrt{300}] / [\sigma^* \times \sqrt{(2 + i/j + j/i)}],$$

with $i, j =$ the number of evaluable subjects in the Placebo and \blacksquare arms at interim, respectively;

$$2.4 \quad C_P = 1 - \Phi((z_{0.65} - Z) / \sqrt{f}), \text{ with } f = 1 - (i+j)/300;$$

$$3.1 \quad \Delta' = \Delta'(N) = 3 \times \sigma^* \times (z_{0.025} + z_{0.10}) / \sqrt{(2N)};$$

$$3.2 \quad \delta'^* = \bar{x}_{\blacksquare} - (\bar{x}_{\blacksquare} + \bar{x}_{Placebo}) / 2;$$

$$3.3 \quad Z' = [(\delta'^* - \Delta') \times \sqrt{450}] / [\sigma^* \times \sqrt{(1.5 + (i+j)/k + (j+k)/4i + (i+k)/4j)}],$$

with $i, j, k =$ the number of evaluable subjects in the Placebo, \blacksquare arms at interim, respectively;

$$3.4 \quad C_{P'} = 1 - \Phi((z_{0.50} - Z') / \sqrt{f'}), \text{ with } f' = 1 - (i+j+k)/450;$$

4. If $\min(C_P, C_{P'}) < 0.20$ then the answer to the question ‘Does the interim analysis outcome suggest that the current study should be stopped for futility based on the probability of a maximum tolerable Phase III non-inferiority sample size of, in total, N subjects?’ is ‘Yes’;
Else the answer is ‘No’.

The analysis results to be communicated (via secure email exchange) exclusively to key stakeholders within the sponsor organization, that is, the Vice President of OTC R&D and the Vice President of Global Medical Affairs, will be presented in – and restricted to – the following table:

<i>Max Tolerable Phase III Sample Size</i>	<i>Interim Analysis Suggests To Stop Study for Futility?</i>
<i>1800</i>	<i>Yes/No</i>
<i>2700</i>	<i>Yes/No</i>
<i>3600</i>	<i>Yes/No</i>

9 REFERENCES

[1] K.K. Gordon Lan and Janet Wittes: *The B-value: A Tool for Monitoring Data. Biometrics, Vol. 44, No. 2, 1988 (pp. 579-585).*