STATISTICAL ANALYSIS PLAN

A 26 week extension study to determine the biological and functional changes in healthy smokers who switched from conventional cigarettes (CC) to Tobacco Heating System 2.2 (THS 2.2) compared to those who continued to smoke CC in the ZRHR-ERS-09-US study

Study Product: Tobacco Heating System 2.2

Sponsor Reference No.: ZRHR-ERS-09-EXT-US

Sponsor:
Philip Morris Products S.A.
Quai Jeanrenaud 5
2000 Neuchâtel
Switzerland

1 STATISTICAL ANALYSIS PLAN APPROVAL SIGNATURES

By signing this page the Statistical Analysis Plan (SAP) is considered final, the signatories agree to the statistical analyses to be performed for this study, and to the basic format of the tables, figures, and listings (TFLs).

Sponsor approval:

[redacted]

11 DEC 2017
Date

Sr. Scientist - Statistics
Philip Morris Products S.A.

[redacted]

12 DEC 2017
Date

Scientist - Statistics
Philip Morris Products S.A.

[redacted]

11-DEC-2017
Date

Clinical Scientist
Philip Morris Products S.A.

[redacted]

12-Dec-2017
Date

Medical Safety Officer
Philip Morris Products S.A.

[redacted]**approval:**

[redacted]

11 Dec 2017

Date

CRO Statistician

2 TABLE OF CONTENTS

3 INTRODUCTION

This SAP has been developed to supplement the statistical analysis described in the clinical study protocol (Final Version 3.0 dated 23 Aug 2016).

This SAP describes the methodology and considerations of the planned analyses of the data collected in both ZRHR-ERS-09-US and ZRHR-ERS-09-EXT-US studies, and lists the TFLs for this study. A detailed description of the TFLs will be provided in a separate TFLs shell document. Any changes to the TFL shell numbering or to a title of a TFL will not require an amendment to this SAP.

This SAP and any amendments will be finalized prior to the lock of the clinical database of the ZRHR-ERS-09-EXT-US study. Any changes to the analyses described in this document or additional analyses performed to supplement the planned analyses, will be documented and described in the clinical study report (CSR).

The preparation of this SAP is based on the following documents:

- International Conference on Harmonisation (ICH) E9 guideline entitled, "Guidance for Industry: Statistical Principles for Clinical Trials" [1]
- ICH E3 guideline entitled, "Guidance for Industry: Structure and Content of Clinical Study Reports" [2]
- Electronic case report forms (eCRF) final version 3.0 (dated 26 Oct 2015).

3.1 Revision History

Version	Date of Revision	Revision
2.0	08-Dec-2017	○ Updated Table 3 and footnotes to clarify covariates used ○ Updated Sections 9, 12.7.1, 12.7.2, and 12.7.3 to indicate that results from inferential analyses will be presented at Months 3 and 6 also ○ Updated Section 9 to indicate that p-values will not be calculated in the analysis of need to cough ○ Updated Section 10.2 to clarify that analyses will be conducted using time-varying product use categories ○ Clarified in Section 10.4 that safety analyses will be performed on 12-month product use categories ○ Updated Section 11.1 to clarify how major PDs are handled ○ Added text to Section 12.1.4 to clarify handling of parameters with values below LLOQ in the denominator of the ratio ○ Clarified in Section 12.3 that analyses will be presented for 6-month and/or 12-month product use pattern categories ○ Clarified in Section 12.7.1.2 that the evaluated

Version	Date of Revision	Revision
		covariates to be included in the model could come from the significant variables found in either the 6-month or 12-month product use categories o Added Section 12.7.1.3.6 to describe multiple imputation sensitivity analysis o Added Section 12.7.1.3.7 to describe new sensitivity analysis for derived parameters with ratio having value below LLOQ o Modified in Section 12.7.2.1 the distribution of product use category levels o Added new listing and sensitivity analysis in Section 12.7.2.2 o Added sensitivity analysis in Section 12.7.2.3 o Added sensitivity analysis in Section 12.7.2.4.1 o Clarified in Section 12.7.2.4.1 that FEV_1 %pred (post-bronchodilator) will not be analyzed with the lung function data since it is already analyzed as a primary endpoint o Removed reference to p-values in Section 12.7.2.5 o Clarified in Section 12.7.3.2 that propensity score model is based on 12-month THS-use pattern category in the THS 2.2 arm o Added SES sub-scores as variables to evaluate in the propensity score model and clarified the terms to include for intent to use questionnaire items in Section 12.7.3.2 o Clarified in Section 12.7.3.2 that the MMRM model will include the same adjustment covariates as used in the primary analysis as well as include the interaction of visit and the PS value fit as a natural cubic spline o Clarified in Section 13.1 that some datasets and TFLs are produced prior to database lock for programming QC purposes o Clarified in Section 16.3.2 that subjects who do not enroll in the extension study will be assigned their 6-month product use category for their 12-month product use category o Updated Section 16.4 table and figure titles to incorporate changes noted above
1.0	25-Apr-2017	Final version

4 ABBREVIATION OF TERMS AND DEFINITIONS OF SPECIAL TERMS

The following abbreviations are used within this SAP.

8-epi-$PGF_{2\alpha}$	8-epi-prostaglandin F2α
11-DTX-B_2	11-dehydro-thromboxane B2
AE/SAE	Adverse Event/ Serious Adverse Event
ALT	Alanine Aminotransferase
Apo A1	Apolipoprotein A1
Apo B	Apolipoprotein B
AP	Alkaline phosphatase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic and Chemical
BP	Blood pressure
BMI	Body Mass Index
BoExp	Biomarkers of exposure
CC	Conventional Cigarettes
CH	Cigarette holder
CI	Confidence Interval
CO	Carbon Monoxide
COHb	Carboxyhaemoglobin
CRO	Contract research organization
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events and Common Toxicity Criteria
CV	Coefficient of Variation
CVD	Cardiovascular disease
ECG	Electrocardiogram
eCRF	Electronic case report forms
EDC	Electronic data capture

EMA	European Medical Association
EOS	End of Study
FAS-AR	Full Analysis Set – as Randomized
FAS-EX	Full Analysis Set – as Exposed
FEF	Forced expiratory flow
FEV_1	Forced Expiratory Volume in 1 second
FRC	Forced residual capacity
FTND	Fagerström Test for Nicotine Dependence
FVC	Forced Vital Capacity
GCP	Good clinical practice
HbA1c	Hemoglobin glycosylated A1c
HDL	High density lipoprotein
HIV	Human Immunodeficiency Virus
HPHCs	Harmful and potentially harmful constituents
hs-CRP	High sensitive C-reactive protein
IC	Inspiratory capacity
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
ITT	Intention-to-treat
ITUQ	Intent to Use THS 2.2 Questionnaire
IXRS	Interactive Web Response System
LDL-C	Low density lipoprotein cholesterol
LLOQ	Lower Limit of Quantification
LS	Least Squares
CC	Conventional Cigarette
MCEQ	Modified Cigarette Evaluation Questionnaire
MedDRA	Medical Dictionary for Regulatory Activities
MPO	Myloperoxidase
MMRM	Mixed Model for Repeated Measurements

Neq	Nicotine equivalents
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNN	Total N-nitrosonornicotine
PI	Principle Investigator
PP	Per-protocol
PS	Propensity score
PT	Preferred Term
QTcB	QT Interval Corrected using Bazett's Formula
QTcF	QT Interval Corrected using Fridericia's Formula
RBC	Red blood cells
RV	Residual Volume
SAP	Statistical Analysis Plan
SD	Standard Deviation
SDTM	Study Data Tabulation Model
SES	Socio-economic status
sICAM-1	Soluble inter-cellular adhesion molecule-1
SOC	System Organ Class
SOP	Standard Operating Procedure
TFL	Tables, Figures, and Listings
THS	Tobacco Heating System
THSts	THS Tobacco Stick
TLC	Total lung capacity
ULOQ	Upper Limit of Quantification
VAS	Visual Analog Scale
VC	Vital capacity
WBC	White blood cell
WHO-DDE	World Health Organization-Drug Dictionary Enhanced

4.1 Definitions for Statistical Analysis

4.1.1 Definition of Special Terms

The following special terms are used in this SAP:

Baseline Value	The last assessment prior to randomization and prior to reported THS use during Run-in.
Conventional Cigarette (CC)	The term 'Conventional Cigarette' refers to commercially available cigarettes (manufactured and hand-rolled) and excludes cigars, pipes, bidis, and other nicotine-containing products.
End of study (individual)	Date of the last visit of the subject plus 28 days of Safety Follow-up Period. For subjects lost to follow-up the individual end of study is the discontinuation date.
End of study (whole study)	End of Study for whole study is the last subject individual EOS.
Enrollment	Subjects will be enrolled at V3 in the main study, after all applicable inclusion and exclusion criteria have been satisfactorily met, and then will be enrolled at V10 in the extension study, after all applicable inclusion and exclusion criteria have been satisfactorily met.
12-month Randomized exposure period	From the check-out on V4 of the main study until the check-out on V16 or earlier discontinuation date after randomization.
Extended exposure period	From the check-out on V10 until the check-out on V16 or earlier discontinuation date after randomization.

Safety follow-up	In the main study, after the check-out on V10 if the subject does not enroll in the extension study or after discontinuation, a 28-day safety follow-up will be done. In the extension study, after the check-out on V16 or after discontinuation, a 28-day safety follow-up will be done.
Screen failure	Subjects who signed the informed consent form (ICF) at V10 but do not meet the entry criteria.
Tobacco Heating System 2.2 (THS 2.2)	THS 2.2 is comprised of the following components: Tobacco Stick, Holder, Charger, a Cleaning Tool, a main power supply, and a USB cable.

4.1.2 Definition of Product Use Pattern Categories

Product use pattern categories within an analysis period will be calculated based on the percentage of THS 2.2 product use of the overall product use, including THS Tobacco Sticks and Conventional Cigarettes (CC). During the analysis period, the percentage of THS 2.2 product use will be calculated as:

$$= 100 \times \frac{Number\ of\ THS\ Tobacco\ Sticks}{Number\ of\ THS\ Tobacco\ Sticks + Number\ of\ CC}$$

General product use pattern categories which will be used for the analysis and presentation of the primary endpoints and for safety summaries are listed in Table 1.

Table 1: Actual Product Use General Pattern Categories

Category Label	General Description
THS-use	• ≥ 1 THSts or CC and • ≥ 70% THSts use over the entire analysis period and • ≥ 70% THSts use on ≥ 50% of the days in the analysis period

Table 1: Actual Product Use General Pattern Categories

Category Label	General Description
Dual-use	• ≥ 1 THSts or CC and • 1% ≤ THSts < 70% over the entire analysis period or • THS-use and CC-use don't apply due to < 50% of the days
CC-use	• ≥ 1 THSts or CC and • < 1% THSts over entire analysis period and • < 1% THSts on ≥ 50% of the days in analysis period
Other-use	General category encompassing subjects with missing product use, or subjects using e-cigarettes or other tobacco products, subjects who quit, or subjects who switched across different use patterns between consecutive analysis periods.

Details of the product use categories are reported in Appendix 16.3 "Data Handling and Derivation of Product Use Pattern Categories". THSts = THS Tobacco Stick; THS=Tobacco Heating System; CC= Conventional Cigarette.

Full details of the product use categories derivation are reported in Appendix 16.3 "Data Handling and Derivation of Product Use Pattern Categories".

General product use categories will be used for both safety summaries and non-safety analysis. More granular product exposure categories will be used for the detailed description of the product use patterns observed in the study.

4.1.3 Categorical Variables

Categorical variables used in this study are shown below (Table 2).

Table 2: Categorical Variables Definitions

Variable	Categories	
BMI (kg/m^2)	Underweight:	<18.5
	Normal range:	≥ 18.5 and < 25.0
	Overweight:	≥ 25.0 and < 30.0
	Obese:	≥ 30.0

Table 2: Categorical Variables Definitions

Variable	Categories	
COPD staging [3] (GOLD1-4 for subjects with FEV_1 [L] / FVC [L] < 0.7; FEV_1 and FVC post-bronchodilator)	Normal:	FEV_1 [L] / FVC [L] ≥ 0.7
	GOLD1: Mild	FEV_1 ≥ 80 %pred
	GOLD2: Moderate	50 ≤ FEV_1 < 80 %pred
	GOLD3: Severe	30 ≤ FEV_1 < 50 %pred
	GOLD4: Very Severe	FEV_1 < 30 %pred
COHb level	≤ 2%	
	> 2%	
Average daily CC consumption level at V1 (based on item 3 of the Smoking History questions)	10-19 cig/day	
	>19 cig/day	
Analysis periods for product use	Period 1 Product Use	]V4, V7]
	Period 2 Product Use	]V7, V10[
	Period 3 Product Use	[V10, V13]
	Period 4 Product Use	]V13, V16[
	Overall Product Use	]V4, V16[
Race	White	
	Black or African American	
	American Indian or Alaska Native	
	Asian	
	Native Hawaiian or Other Pacific Islander	
	Any combination from eCRF e.g. White and Black	
Ethnicity	Hispanic or Latino	
	Not Hispanic or Latino	
Race/Ethnicity	White Hispanic	
	White not Hispanic	
	Black or African American	
	Asian	
	Other	
Caucasian origin	Caucasian:	White, and not Hispanic or Latino
	Non-caucasian:	Not white, or Hispanic or Latino

Table 2: Categorical Variables Definitions

Variable	Categories	
Labels for analysis timepoint	Screening	V1
	Enrollment	V3
	Day 1	V4
	Month 1	V5
	Month 2	V6
	Month 3	V7
	Month 4	V8
	Month 5	V9
	Month 6	V10
	Month 7	V11
	Month 8	V12
	Month 9	V13
	Month 10	V14
	Month 11	V15
	Month 12	V16

4.1.4 Covariates for Smokers' Health Profile Analysis

The variables listed as "defined covariates" in Table 3 will be included in the analysis model. The "evaluated covariates" specified in Table 3 will be considered for their inclusion in the analysis model, as described in Section 12.7.1.2 Smokers' Health Profile Endpoints Evaluation".

Table 3: Baseline Covariates for "Smokers' Health Profile" Analysis

Endpoint	Defined Covariates[1]	Evaluated Covariates[1]
HDL-C	Age, Smoking intensity	Smoking duration, Diet, Alcohol intake, Exercise, BMI
WBC	Age, Smoking intensity	Smoking duration, Race/Ethnicity, Sleep deficit
sICAM-1	Age, Smoking intensity	Smoking duration
11-DTX-B_2	Age, Smoking Intensity	Other exposure
8-epi-PGF2α	Age, Smoking intensity	Smoking duration, BMI, Weight
COHb	Age, Smoking intensity	Other exposure

Table 3: Baseline Covariates for "Smokers' Health Profile" Analysis

Endpoint	Defined Covariates[1]	Evaluated Covariates[1]
FEV_1 %pred	Smoking intensity	Age[2], Race/Ethnicity[2] , Height[2], Diet, Exercise, BMI, Weight, Smoking duration
Total NNAL	Age, Smoking intensity	Other exposure

1 Listed covariates are defined as follows:

- Smoking intensity: average daily CC consumption measured by item 3 from the smoking history questions

- Smoking duration: measured by item 2 from the smoking history questions

- Age, Weight, Height, BMI : continuous variable

- Race/Ethnicity: (White Hispanic, White not Hispanic, Black or African American, Asian, Other) in substitution of the Caucasian origin variable if found significant between THS-use and CC-use product use exposure groups at]V4, V16[

- Diet, Alcohol intake, Exercise (total weekly time), sleep deficit, and Other exposure as collected in the Lifestyle questions

2 These variables are accounted for in the %predicted assessment. Race/Ethnicity will be included in the statistical model in place of Caucasian origin variable if found significant between THS-use and CC-use product use exposure groups at]V4, V16[.

5 STUDY OBJECTIVES AND ENDPOINTS

5.1 Primary Objectives and Endpoints

1. To determine the changes to the selected clinical risk endpoints ("smokers' health profile") in smokers who have switched from CC to THS 2.2 as compared to those who continued to smoke CC.

   Endpoints in the "smokers' health profile" measured at V16 (Week 52):

   - High density lipoprotein cholesterol (HDL-C) in serum.
   - White blood cell total count (WBC) in blood.
   - Soluble Intercellular adhesion Molecule 1 (sICAM-1) in serum.
   - 11-dehydrothromboxane B2 (11-DTX-B_2) in urine (expressed as concentration adjusted for creatinine).
   - 8-epi-prostaglandin $F_{2\alpha}$ (8-epi-$PGF_{2\alpha}$) in urine (expressed as concentration adjusted for creatinine).
   - Carboxyhemoglobin (COHb) in blood.
   - Forced expiratory volume in 1 second (FEV_1 post-bronchodilator, expressed as % predicted [FEV_1 %pred]).
   - Total 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (Total NNAL) in urine (expressed as concentration adjusted for creatinine).

5.2 Secondary Objectives and Endpoints

2. To evaluate self-reported product use (CC and THS 2.2) over the duration of the study.

   Endpoint (measured daily):

   - Number of CC or THS Tobacco Sticks used daily as reported in the self-reported product use electronic diary.

3. To determine the reduction in exposure to harmful and potentially harmful constituents (HPHCs) in smokers who have switched from CC to THS 2.2 as compared to those who continued to smoke CC.

   Endpoints at V16 (Week 52):

   - Biomaker of exposure (BoExp) to carbon monoxide (CO): CO in exhaled breath (expressed as ppm).
   - BoExp to N-nitrosonornicotine: total N-nitrosonornicotine (total NNN) in urine (expressed as concentration adjusted for creatinine).

4. To determine the levels of nicotine exposure in smokers who have switched from CC to THS 2.2 as compared to those who have continued to smoke CC.

   Endpoints (BoExp to nicotine) over the duration of the study:

- Nicotine equivalent (Neq): molar sum of free nicotine, nicotine-glucuronide, free cotinine, cotinine-glucuronide, free *trans*-3'-hydroxycotinine, *trans*-3'-hydroxycotinine-glucuronide in urine (expressed as concentration adjusted for creatinine).
- Nicotine and cotinine in plasma expressed in ng/mL.

5. To describe the changes of biological or functional markers associated with respiratory diseases and cardiovascular diseases (CVD) in smokers who have switched from CC to THS 2.2 as compared to those who have continued to smoke CC.

   Endpoints associated with respiratory diseases measured at V16 (Week 52):

   - Lung function:
     - Spirometry pre- and post-bronchodilator: FEV_1, forced vital capacity (FVC), FEV_1/FVC, forced expiratory flow (FEF 25-75) and bronchodilator reversibility in FEV_1.
     - Lung volume pre-bronchodilator: forced residual capacity (FRC), vital capacity (VC), total lung capacity (TLC), and inspiratory capacity (IC).
   - Cough symptoms (intensity and frequency), amount of sputum production, and bothersomeness of cough symptom from the cough questionnaire.

   Endpoints associated with CVD measured at V16 (Week 52):

   - Myeloperoxidase (MPO), apolipoprotein A1 and B (Apo A1 and Apo B), high sensitivity C-reactive protein (hs-CRP), and low density lipoprotein cholesterol (LDL-C), in serum.
   - Fibrinogen and homocysteine in plasma.
   - Albumin in urine (expressed as concentration adjusted to creatinine).
   - Platelet count and hemoglobin glycosylated (HbA1c in whole blood).
   - Blood pressure, weight, and waist circumference.

6. To describe the changes in subjective effects of smoking in smokers who have switched from CC to THS 2.2 as compared to those who have continued to smoke CC.

   Endpoint over the duration of the study:

   - Product evaluation: subscales from the modified cigarette evaluation questionnaire (MCEQ).

7. To describe the intention to use THS 2.2 in smokers who have switched from CC to THS 2.2.

   Endpoint at V16 (Week 52):

- Intention to use associated with THS 2.2: item scores from intent to use questionnaire for THS 2.2 (ITUQ) only for subjects in the THS 2.2 arm.

8. To describe the change in tobacco dependence in smokers switching from CC to THS 2.2 as compared to those continuing to smoke CC.

   Endpoints at V16 (Week 52):

   - Score from the Fagerström (FTND) questionnaire.
   - Time to first cigarette from the FTND questionnaire.

9. To describe the effect of combined product use (Dual-use) over the study on the components of the "smokers' health profile".
   Endpoint:
   - The levels of the components in the "smokers' health profile" and the number of CC and THS Tobacco Sticks used daily as reported on the self-reported product use electronic diary.

10. To evaluate the safety profiles associated with THS 2.2 and CC.
    Endpoints over the duration of the study:
    - Adverse events (AEs), serious adverse events (SAEs) and device events, including THS 2.2 malfunction/misuse.
    - Vital signs.
    - Electrocardiogram (ECG).
    - Clinical chemistry, hematology and urine analysis safety panel.
    - Physical examination.
    - Concomitant medications.

5.3 Additional Endpoints

- Urine pregnancy.
- Attempts to quit smoking.
- Residual volume.
- Apo B/Apo A1 ratio.
- Albumin quantity excreted in 24h.

5.4 Study Hypotheses And Evaluation Criteria

5.4.1 Hypotheses

This study has no formal pre-specified hypotheses associated with the study objectives. Estimates of THS 2.2 effect will be presented and 95% CI will accompany all effect estimates. Two-sided p-values (null hypothesis of no difference between groups) will be presented for baseline comparability and for the THS 2.2 effect on endpoints in the "smokers' health profile", on BoExp to HPHCs, and on BoExp to nicotine. The objective of the study is to determine the effect of THS 2.2 compared to CC at Week 52 on the components of the "smokers' health profile" as defined by PMI (Protocol Table S1), and to provide additional information to the results of the original study (ZRHR-ERS-09-US) for a prolonged exposure period.

5.4.2 Evaluation Criteria

The study will target to describe the 95% confidence intervals (CI) of the THS 2.2 effect as compared to CC at Week 52 on the components of the "smokers' health profile" estimated with a precision of ±75% of the anticipated THS 2.2 effect at week 52.

6 INVESTIGATIONAL PLAN

6.1 Study Design

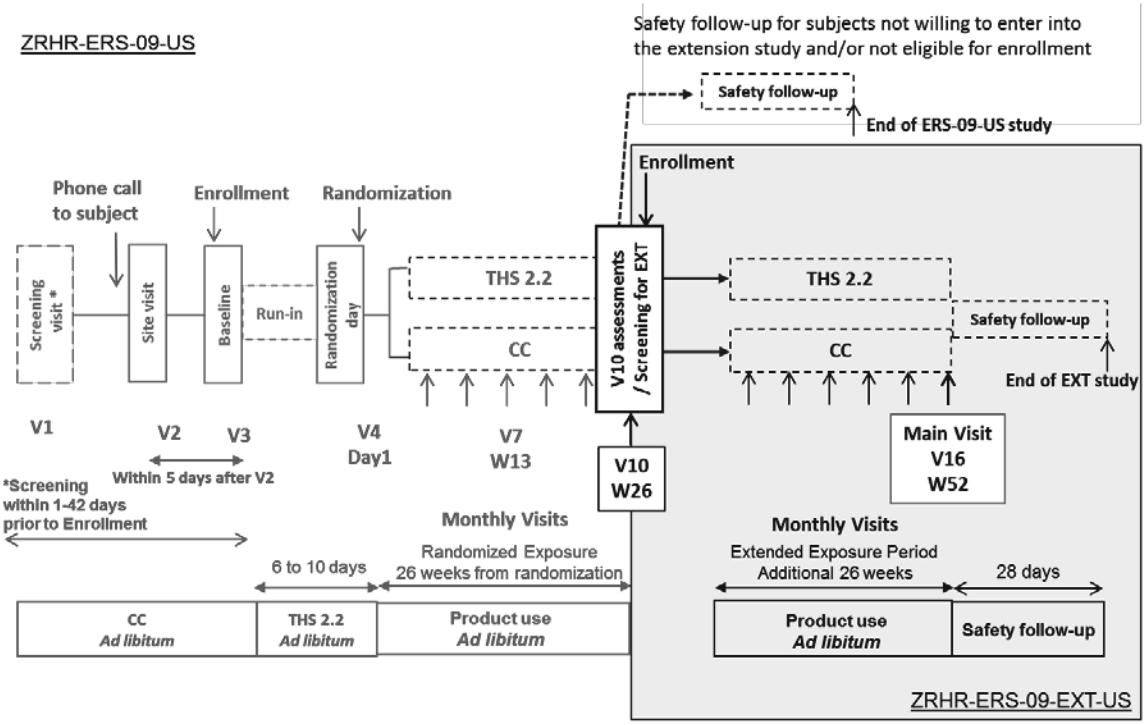
This study is a 26 week extension study of the original study (ZRHR-ERS-09-US), a randomized, controlled, open-label, 2-arm, parallel group study design conducted over a 26 week period. This 6-month extension study will be conducted as a separate study, as a follow-up of the 6-month randomized exposure period of the original study, extending the exposure to a 12-month randomized exposure period from V4-V10 (Day 1 - Week 26, 6 months) to V4-V16 (Day 1 - Week 52, 12 months), using the same sites. Subjects will continue to use the product they were randomized to in the original study (THS 2.2 arm or CC arm).

The extension study will be an *ad libitum* smoking study with unrestricted product use for the duration of the study (including during site visits).

The end of the study (EOS) for the individual is defined as the date of the last visit of the subject plus 28 days of Safety Follow-up Period. For subjects lost to follow-up the individual end of study is the discontinuation date. End of Study for whole study is the last subject individual EOS.

Subjects that terminate the study after enrollment and prior to completion of the extension study will undertake early termination procedures.

Figure 1 Study Flowchart



Abbreviations: CC = conventional cigarette(s); THS = Tobacco Heating System

6.2 Selection of Study Population

All subjects who have completed the 26 week randomized exposure period (V10) of the original study will be potentially eligible for enrollment into the extension study, depending on whether or not they want to participate and if they meet the entry criteria.

The maximum number of subjects that can be enrolled is pre-defined by the number of subjects completing the 26 week period of the original study. The target population in the original study was approximately 950 female or male, healthy adult smokers. All races/ethnicities were considered eligible, although a quota was applied to subjects of

Caucasian origin (Race is White and Ethnicity is Not Hispanic or Latino), as they should not represent more than 75% of the randomized subjects.

6.2.1 Inclusion Criteria

At V10, each subject must meet the following criteria to be eligible to enter in the extension study:

Inclusion Criteria	**Screening at V10**
1. Subject has completed V10 of the original study (ZRHR-ERS-09-US).	X
2. Subject has signed the ICF at V10 and is able to understand the information provided in the ICF.	X
3. The subject is willing to comply to study procedures and to continue to use the product he/she was allocated to during the original study (THS 2.2 or CC) for an additional 26 weeks.	X

6.2.2 Exclusion Criteria

Subjects who meet any of the following exclusion criteria must not be enrolled into the study:

Exclusion Criteria	**Screening at V10**
1. As per judgment of the Principle Investigators (PIs) or designee(s), the subject cannot participate in the study for any reason (e.g. medical, psychiatric and/or social reason).	X
2. Any medical conditions that in the opinion of the PI(s) or designee(s) would jeopardize the safety of the participant.	X
3. Subject has made a quit attempt* during the original study.	X
4. For women only: Subject is pregnant (does not have a negative pregnancy test at V10) or is breast feeding.	X

5. For women only : Subject does not agree to use an acceptable method of effective contraception**.	X

* From using tobacco-containing products (e.g. THS 2.2 and CC). This information will be captured from the original study.

** *e.g.*, Intrauterine device, intrauterine system, established use of oral/injectable/implantable/transdermal hormonal methods, barrier methods of contraception (condoms, occlusive caps) with spermicidal foam/gel/film/suppository, vasectomized partner(s), or true abstinence (periodic abstinence and withdrawal are not effective methods) from V10 until V16. Hysterectomy, tubal ligation, bilateral oophorectomy or post-menopausal status are reasons for not needing to use birth control. Post-menopausal status is defined as women who have not experienced menses for greater than 12 months.

6.3 Product Allocation and Blinding

6.3.1 Methods of Assigning Subjects to Product Arms

During the extension study, subjects will remain in the study arm they were randomized to in the original study (THS 2.2 arm or CC arm in a 1:1 ratio).

Randomization was done during the original study at V4 through the interactive voice and web response system (IXRS).

6.3.2 Blinding

This is an open-label study; therefore, the subjects and Investigators will be unblinded to subject's arm after randomization at V4 during the original study. However, and similar to the original study there will be a limited degree of blinding during the conduct of the extension study, including the data review and data analysis process. In particular, some members of PMI and CRO personnel will be blinded as summarized in Table 4.

Table 4: Blinding Scheme

Blinded Study Personnel	Blinded Data	End of Blinding Period
PMI and [redacted] study statisticians	Actual values of primary endpoints at V16[1]	After the SAP finalization or database lock, whichever comes last.
PMI clinical scientist	Actual values of primary endpoints at V16[1]	After the finalization of PMI blind database review[2]. Can be actively unblinded when appropriate.

Table 4: Blinding Scheme

Blinded Study Personnel	Blinded Data	End of Blinding Period

[1] To avoid indirect unblinding of the "actual values of primary endpoints" additional data were blinded in the eCRF (e.g., adverse event terms). After SAP finalization or database lock of the original study, whichever comes last, the whole study team will be unblinded to all data collected in the original study, following the process described in the original study protocol.

[2] As part of the PMI Quality Control activity, data listings will be reviewed by [redacted] and PMI before database lock. Full details will be available in the data review plan.

Any PMI and [redacted] personnel who are not listed in the above table will be unblinded by default.

Unblinded information will not be shared with the blinded study team, until the end of the blinding period (Table 4). PMI will receive blinded and unblinded data for the pre-analysis data review as planned in the data review plan. Blinded data will be accessible by the blinded study personnel in a masked format or presented independent of the subject identifier so to ensure that data cannot be associated within or to a subject. Unblinded data will only be reviewed by the unblinded study team.

6.3.3 Adherence to Product Allocation

The number of CC, THS Tobacco Sticks, and other nicotine/tobacco containing product used daily, as reported by the subject and collected in the electronic diary, will be used to monitor product use and evaluate adherence to product allocation.

Adherence to THS 2.2 is defined as at least 70% THS 2.2 use with no more than 5 CC during each single day of the randomized exposure period of the main study. The same criteria will be adopted for the extended 6-month period. The choice of 5 CC per day is based on the definition of the smoking cessation endpoint: abstinence with occasional slips allowed, as reported in European Medical Association (EMA) guidelines [4].

Subjects in the THS 2.2 and CC arm will be considered adherent to the randomized product only if they belong to THS-use and CC-use product categories, respectively (see Section 4.1.2 "Definition of Product Use Pattern Categories").

For both the THS 2.2 and CC arms, any subjects reporting to having quit using tobacco-containing products (*e.g.* THS 2.2 and CC) will be considered not adherent to the study product allocated.

Diary adherence is defined as at least 75% diary use during the randomized exposure period and no more than 7 consecutive days of missing diary use during the randomized exposure period.

Product adherence and diary adherence will be used for the definition of the Per Protocol population (see Section 10.3 “Per Protocol ”).

7 DERIVED AND COMPUTED VARIABLES

Mean change from Baseline (Baseline is defined in Section 4.1.1 "Definition of Special Terms") is the mean of all individual subjects' change from Baseline values. Each individual change from Baseline will be calculated by subtracting the individual subject's Baseline value from the value at the timepoint. The individual subject's change from Baseline values will be used to calculate the mean change from Baseline.

Mean percent change from Baseline is the mean of all individual subjects' percent change from Baseline values. Each percent change from Baseline will be calculated by subtracting the individual subject's Baseline value from the value at the desired timepoint and then dividing this calculated value by the individual subject's Baseline value and multiplying by 100. When the Baseline value is 0 for at least 1 subject, the percentage change from Baseline will not be calculated and the number of such cases will be tabulated as "Baseline=0" in any descriptive summaries.

Mean percent relative change from Baseline is estimated as: 100*(exp(log_mean_change)-1) where log_mean_change is computed by averaging differences: log(post baseline value)-log(baseline value) at the desired timepoint. When the Baseline value is 0 for at least 1 subject, the percent relative change from Baseline will not be calculated and the number of such cases will be tabulated as "Baseline=0" in the descriptive summaries.

Caucasian origin variable will be derived to be considered in the analysis. Caucasian origin is defined as 'Caucasian' (Race is White and Ethnicity is Not Hispanic or Latino) or 'Non-caucasian' (Race is not White or Ethnicity is Hispanic or Latino).

Reported BMI will be calculated at site from the body weight and height using the following formula:

$$\text{BMI} = \frac{\text{weight in kilograms}}{\text{height in meters}^2} \quad (\text{kg/m}^2)$$

The QT interval corrected using Bazett's formula (QTcB) will be calculated as follows:

$$QTcB = \frac{QT}{\sqrt[2]{\left(60/HR\right)}}$$

The QT interval corrected using Fridericia's formula (QTcF) will be calculated as follows:

$$QTcF = \frac{QT}{\sqrt[3]{\left(60/HR\right)}}$$

The geometric coefficient of variation (CV) will be calculated using the following formula:

$$CV = 100\sqrt{e^{\text{var}} - 1}$$

where var = the variance from the log-transformed data.

7.1 Clinical Risk Endpoints and Biomarkers of Exposure

7.1.1 Clinical Risk Endpoints and Biomarkers of Exposure in Urine

The adjustment for creatinine of the urinary clinical risk endpoints and biomarkers will be calculated as:

$$\text{Clinical Risk Endpoint or Biomarker (adjusted for creatinine)} = \frac{[\text{Clinical Risk Endpoint or Biomarker}]}{[\text{Creatinine}]}$$

where the [ ] indicated concentrations measured from the same 24 hour urine collection.

7.1.2 Nicotine Equivalents

The concentration of Neq in 24 hours will be derived according to the formula below. The concentrations reported for free nicotine and its five major metabolites will not be used as analysis variables.

Neq [mg/L] = (free nicotine[µmol/L] + nicotine-glucuronide [µmol/L]
+ free cotinine [µmol/L] + cotinine-glucuronide [µmol/L]
+ free trans-3'-hydroxycotinine [µmol/L]
+ trans-3'-hydroxycotinine-glucuronide [µmol/L])
*162.2[µg/µmol] / 1000 [µg/mg]

N.B. All concentrations must be in µmol/L before applying the above formula.

The conversion factors will be applied as follows:

Free nicotine	The molecular weight is 162.232 g/mol [5]. Therefore to convert nicotine from ng/mL to nmol/L, the result in ng/mL is multiplied by 6.164.
Nicotine glucuronide	The molecular weight is 338.356 g/mol [6]. Therefore to convert nicotine from ng/mL to nmol/L, the result in ng/mL is multiplied by 2.955.
Cotinine	The molecular weight is 176.218 g/mol [7]. Therefore to transform cotinine from ng/mL to nmol/L, the result in ng/mL will be multiplied by 5.675.
Cotinine-glucuronide	The molecular weight is 352.341 g/mol [8]. Therefore to transform cotinine from ng/mL to nmol/L, the result in ng/mL will be multiplied by 2.838.
Trans-3'hydroxycotinine	The molecular weight is 192.217 g/mol [9]. Therefore to transform trans-3' hydroxycotinine from ng/mL to nmol/L, the result in ng/mL is multiplied by 5.202.
Trans-3'hydroxycotinine-glucuronide	The molecular weight is 368.34 g/mol [10]. Therefore to transform trans-3' hydroxycotinine from ng/mL to nmol/L, the result in ng/mL is multiplied by 2.715.

7.2 Questionnaires

7.2.1 Fagerström Test for Nicotine Dependence (FTND)

The FTND [12] will be used in its revised version, as updated in 2012 [13]. These questions are to be answered by the subjects themselves. It is conducted at V1 to determine subject's dependence on nicotine and at V10 and V16 to assess any changes in their dependence on nicotine.

Table 5 describes the six questions the questionnaire consists of, and the scores associated with each question.

The FTND total score will be derived by summing the individual item scores if all items are non-missing; if any items are missing the total score will be set to missing. For the FTND total score, descriptive statistics and frequency tables according to the following classification will be provided [13]:

Mild	0 – 3
Moderate	4 – 6
Severe	7 – 10

Table 5: Scoring for the Fagerström Test for Nicotine Dependence

	FTND Question	Response	Score
1	How soon after you wake up do you smoke your first cigarette?	Within 5 minutes	3
		6 to 30 minutes	2
		31 to 60 minutes	1
		After 60 minutes	0
2	Do you find it difficult to refrain from smoking in places where it is prohibited?	Yes	1
		No	0
3	Which cigarette would you most hate to give up?	The first one in the morning	1
		Any other	0
4	How many cigarettes per day do you smoke?	10 or less	0
		11 to 20	1
		21 to 30	2
		31 or more	3
5	Do you smoke more frequently during the first hours after waking up than during the rest of the day?	Yes	1
		No	0
6	Do you smoke even if you are so sick that you are in bed most of the day?	Yes	1
		No	0

7.2.2 Modified Cigarette Evaluation Questionnaire

The MCEQ [14] will be completed by the subject at V3, V7, V10, and V16.

The MCEQ consists of 12 items as presented in Table 6.

Table 6: Modified Cigarette Evaluation Questionnaire - Questions and Subscales

	Question	Subscale
1	Was smoking satisfying?	Smoking Satisfaction
2	Did cigarettes taste good?	Smoking Satisfaction
3	Did you enjoy the sensation in your throat and chest?	Enjoyment of Respiratory Tract Sensations
4	Did smoking calm you down?	Psychological Reward

Table 6: Modified Cigarette Evaluation Questionnaire - Questions and Subscales

	Question	Subscale
5	Did smoking make you feel more awake?	Psychological Reward
6	Did smoking make you feel less irritable?	Psychological Reward
7	Did smoking help you concentrate?	Psychological Reward
8	Did smoking reduce your hunger for food?	Psychological Reward
9	Did smoking make you dizzy?	Aversion
10	Did smoking make you nauseous?	Aversion
11	Did smoking immediately relieve your craving for a cigarette?	Craving Reduction
12	Did you enjoy smoking?	Smoking Satisfaction

Items are assessed on a 7-point scale, ranging from 1 (not at all) to 7 (extremely). Higher scores indicate greater intensity on that scale.

The subscales will be derived by averaging the individual non-missing item scores if at least 50% are non-missing, otherwise the subscale score will be set to missing.

7.2.3 Smoking History Questions

At V1, subjects will be asked 5 questions related to their current and past smoking behavior, shown in Table 7.

Table 7: Questions on Smoking History/Habits

	Question
1	Have you smoked for at least the past 10 years? (Yes/No)
2	How many years have you smoked?
3	On average, how many cigarettes per day have you smoked over the last year
4	On average, how many cigarettes per day have you smoked since you started smoking
5	On average, how would you describe your e-cigarette use over the last year? (Check one: Daily - How much use per day?/ Weekly - How much use per week? / Sporadically [less than once per week] / Tried e-cigarettes. [between 1–10 uses] / Never tried e-cigarettes)

The smoking intensity is defined as the average number of CC smoked during the last year and will be derived from the self-reported subjects' answer to item 3. The smoking history intensity evaluating the amount a person has smoked over his/her life, expressed as number of pack-years, will be derived as: [item 2] x [item 4]/20.

7.2.4 Prochaska 'Stage of Change' Questionnaire: Intention to Quit Smoking

Subjects will complete the Prochaska 'Stage of Change' questionnaire at V1.

The Prochaska 'Stage of Change' questionnaire will be used to assess the smokers' mental state for the intention to quit (Precontemplation, Contemplation, Preparation, Action, and Maintenance). The details of the staging algorithm [15, 16] are provided in the Appendix 5 of the study protocol.

The questions are shown in Table 8.

Table 8: Prochaska 'Stage of Change' Questionnaire

	Question	Responses
1	Are you currently a smoker?	Yes No, I quit within the last 6 months No, I quit more than 6 months ago No, I have never smoked
2	In the last year, how many times have you quit smoking for at least 24 hours?	
3	Are you seriously thinking of quitting smoking?	Yes, within the next 30 days Yes, within the next 6 months No, not thinking of quitting

7.2.5 Intent to Use THS 2.2 Questionnaire

Subjects will be asked about their intention to use THS 2.2 using the Intent to Use THS 2.2 Questionnaire (ITUQ) at V4, V10, and at V16. This questionnaire consists of 7 items as presented in Table 9.

Table 9: Intent To Use Questionnaire

	Question	Responses
1	How likely or unlikely are you to use THS 2.2 regularly?	definitely not to definitely.
2	How likely or unlikely are you to use THS 2.2 and continue to smoke cigarettes? (*)	definitely not to definitely.
3	How likely or unlikely are you to use THS 2.2 and continue to use other tobacco-nicotine products besides cigarettes (such as nicotine patch, pipes, or chewing tobacco)? (*)	definitely not to definitely.
4	How likely or unlikely are you to use THS 2.2 to help you quit cigarettes? (*)	definitely not to definitely.
5	How likely or unlikely are you to use THS 2.2 rather than quitting cigarettes? (*)	definitely not to definitely.
6	How likely or unlikely are you to use THS 2.2 rather than use a product to help quit cigarettes (such as nicotine patches or nicotine gums)? (*)	definitely not to definitely.
7	How soon would you begin to use THS 2.2 regularly? (*)	Never Within a Week Within a Month Within 3 Months Within 6 Months More than 6 Months

* Ask subjects who answer “to use THS 2.2” for item 1 with a different answer than “Definitely not”

Items are assessed on a 6-point scale, ranging from “definitely not” to “definitely” for the first 6 questions. Subjects will answer to items 2 to 7 only if they provide answer different from “definitely not” to item 1. Higher scores indicate greater likelihood to use the THS 2.2 product. Item 7 is rated on an ordinal scale, with 6 timeframe categories. For subjects who enter ‘Definitely not’ to item 1, they will be assigned ‘Definitely not’ to items 2-6 and ‘Never’ to item 7. Responses from items 1 to 6 will also be mapped into categories of ‘Positive intention’ (responses of “very likely” or “definitely”) and ‘Not positive intention’ (responses of “definitely not”, “very unlikely”, “somewhat unlikely”, or “somewhat likely”).

7.2.6 Product Preference Question

At V4, the following question will be asked to subjects prior to randomization (after the ITUQ questionnaire but prior to the question on the readiness of the subject to use the THS 2.2 for 26 weeks): “Which product would you prefer to be randomized to:” with

possible answers: THS 2.2 / CC / No preference. As an instruction to answer this question, subjects will be informed that their response will not influence the randomization process.

7.2.7 Socio-Economic Status Questionnaire

At V3, subjects will complete a series of questions related to their education, occupational status, size and annual income of their household. These data will be used to create measures of educational attainment (Low / Moderate / High) and annual household income (Low / Moderate/ High), as detailed in Table 10 below (see [17]).

Table 10: Classification for the Socio-Economic Status Questionnaire

Question	Answer	Category
Question 1. What is your highest educational level?	1 Less than High School	Low
	2 High School Graduate	Moderate
	3 Some High school or general education development (GED)	Moderate
	4 College Graduate	High
	5 Advanced Degree	High
Question 2. What is your current occupational status?	1 Working now	Not applicable
	2 Only temporarily laid off, sick leave or maternity leave	
	3 Looking for work, unemployed	
	4 Retired	
	5 Disabled, permanently or temporarily	
	6 Homemaker, keep housing	
	7 Student	
	8 Other	
Question 3. How many people are currently living in your household, including yourself?	Number (0-99)	Not applicable
Question 4. Of these people, how many are children under the age of 18?	Number (0-99)	Not applicable
Question 5. Of these people, how many are adults?	Number (0-99)	Not applicable
Question 6. Of the adults, how many bring income into the household?	Number (0-99)	Not applicable

Table 10: Classification for the Socio-Economic Status Questionnaire

Question		Answer	Category
Question 7. Which of these categories best describes your total combined family income for the past 12 months?	1	Less than $10,000	Low
	2	$10,000 to $29,999	Low
	3	$30,000 through $44,999	Moderate
	4	$45,000 through $59,999	Moderate
	5	$60,000 through $74,999	High
	6	$75,000 through $99,999	High
	7	$100,000 through $149,999	High
	8	$150,000 and over	High
	9	I do not know	< Missing>
	10	Prefer not to say	< Missing>

For Question 1 (Education) and Question 7 (Annual household income), if multiple answers are obtained the higher degree or income will be chosen; the subject cannot be classified if the answer is missing.

Education and income categories will be combined to create a composite measure for socio-economic status (SES) with categories defined as follows:

- Low (Low education and Low income categories)
- Moderate (Low income and Moderate education, Low income and High education, Moderate income and Low education, and High income and Low education)
- High (Moderate income and Moderate education, Moderate income and High education, High income and Moderate education, and High income and High education).

Respondents who do not report either income or education will be excluded from the analysis of the composite SES.

7.2.8 Cough Questionnaire

Subjects will self-report the respiratory symptom ‘cough’ using a visual analog scale (VAS), three Likert scale questions, and one open ended question at V3, V7, V10, and V16.

Subjects will be asked if they have experienced a regular need to cough, (*i.e.,* whether they have coughed several times in the previous 24 hours prior to assessment). If the answer is ‘yes’, subjects will be asked to complete the questionnaire.

The VAS will assess how bothersome cough is to the subject ranging from ‘not bothering me at all’ to ‘extremely bothersome’, and this will be given a numeric value between 0 and 100, measured on a 100 mm scale.

Subjects will also be asked to assess the intensity and frequency of cough and the amount of sputum production on Likert scales as presented in Table 11.

Table 11: Cough Assessment Likert Scales

	Question	Likert Scale
1	The intensity of cough	1 = very mild 2 = mild 3 = moderate 4 = severe 5 = very severe
2	The frequency of cough	1 = rarely 2 = sometimes 3 = fairly often 4 = often 5 = almost always
3	The amount of sputum production	0 = no sputum 1 = a moderate amount of sputum; 2 = a larger amount of sputum; 3 = a very large amount of sputum.

7.2.9 Lifestyle Assessment

Subjects will be asked questions to capture baseline covariates to assess lifestyle characteristics at V3. The replies to the questions reported in Table 12 will be considered for their inclusion in the analysis as described in Section 4.1.4 "Covariates for Smokers' Health Profile Analysis".

Table 12: Lifestyle Assessment Questions

	Question	Answers
1	How many times per week do you eat fast food? (*e.g.*, hamburgers, hot dogs, pizza, french fries)	Number (0-99)
2	How many alcoholic drinks do you have a day?	Number (0-99)
3	How many times a week do you exercise?	Number (0-99)
4	What is the average duration of each exercise session? (record in minutes)	Number (0-999)
5	How many nights per week do you sleep less than 6 hours?	Number (0-9)
6	Do you live in a household with one or more smokers, other than yourself?	Yes/No

The total weekly exercise time (minutes/week) will be in general derived as [item 3]x[item 4] of the Lifestyle assessment questions. Subsequent to data review findings, the following derivation rules will also be applied:

- If the result is greater than 3360 minutes (corresponding to an average daily exercise time of 8 hours) then the total weekly exercise time is set to 3360 minutes.

- If the reply to item 3 is missing the total weekly exercise time is missing, apart from when item 4 is zero which results into a total weekly exercise time equal to zero. Similarly, if the reply to item 4 is missing the total weekly exercise time is missing, apart from when item 3 is zero which results into a total weekly exercise time equal to zero.

8 SAMPLE SIZE JUSTIFICATION

The sample size for this study was defined based on the sample size requirements of the original study, for which the sample size was calculated to ensure an overall study power of at least 90%, while maintaining at least 80% power to detect the expected effect of THS 2.2 as compared to CC for each individual endpoint of the "smokers' health profile", over 26 weeks of exposure. This required 475 subjects randomized per group accounting for an anticipated 75% of the THS 2.2 arm to be Mostly THS 2.2 users (i.e. report at least 70% THS 2.2 use).

Based on the assumptions that 30% of the 950 subjects will not be enrolled or will drop out from the extension study and a potential decrease from 75% to 50% of Mostly THS 2.2 users among subjects randomized to THS 2.2, the study will have at least 90% probability to determine the effect of THS 2.2 as compared to CC on FEV1 at V16 with a margin of error (95% CI) of at most ±1.5 %pred. The anticipated SD of 6.4 %pred was estimated using the results of the "Lung Health Study" [1].

Following our current understanding of product use data in the main study, the definition of the product use categories was redefined and the underlying assumptions re-evaluated so that approximately 60% of the subjects in the THS 2.2 arm are expected to be in the THS-use category, with an anticipated 20% of subjects drop-out or reporting insufficient product use data in the two arms of the main study. Based on the 65% rate of enrollment into the extension study observed among subjects randomized in the main study, the extension study will have more than 90% probability to determine the effect of THS-use compared to CC-use on FEV1 at V16 with a margin of error (95% CI) of at most ±1.5 %pred.

9 CHANGES FROM THE PROTOCOL SPECIFIED STATISTICAL ANALYSIS

The following changes from the protocol specified statistical analysis were identified prior to database lock:

- Additional endpoints collected in the study but not listed in Section 3 of the protocol are reported in Section 5.3 "Additional Endpoints". This includes residual volume which was not included in the protocol objectives but will be derived for interpretation of lung volume. Also, the Apo B/Apo A1 ratio will be analyzed and results will be used for the interpretation of Apolipoprotein data. The quantity excreted in 24h will be derived from the laboratory for Albumin as additional risk marker associated with CVD and renal disease. Spirometry (clinical interpretation and COPD categories), Respiratory symptoms (cough assessment VAS and Likert scales), and Body weight and BMI were not detailed as safety endpoints in Section 3 (Study Objectives) of the protocol but were considered for safety analysis in Section 12.6.4 of the protocol and are analyzed accordingly, in line with analyses from the original study.

- The definition of safety periods in Section 12.6.4 of the protocol showed that V3 is included in the Screening period and excluded from the Product trial period. However, in line with emergent events starting at enrollment, V3 is now excluded from the Screening period and included with the Product trial period instead.

- The distribution of product use cross-tabulation summary outlined in Section 12.7.2.1 will be summarized over the entire 12 month exposure period across the main and extension studies rather than the 6 months specified in the protocol.

- Clarified in Section 10.1 which visits are included in FAS-AR definition.

- The analysis of cough symptoms specified in Section 12.6.2 of the protocol was updated to present odds ratios rather than incidence rate differences because of use of the logistic model. The original model using the identity link to provide incidence rate differences could lead to convergence issues and so the approach was changed. In addition, age and smoking intensity were included in the logistic model, for consistency with the model used for other endpoints. Also, p-values will not be calculated.

- Inferential analyses in Sections 12.7.1, 12.7.2, and 12.7.3 will be presented at Months 3 and 6 in addition to Month 12.

- As noted in Section 10.2, the FAS-EX analyses of primary and secondary endpoints will be performed using 6-month product use categories for Month 1 – Month 6 data and 12-month product use-categories for Month 7 – Month 12 data rather than using the 12-month product use categories for all timepoints as planned in the protocol, to allow for consistency with the results in the original study.

10 ANALYSIS POPULATIONS

The Full Analysis Set – as Exposed (FAS-EX) will be the primary analysis set for clinical risk endpoints, BoExp, and questionnaires, based on the overall product use pattern categories. Supportive statistical considerations are reported in Appendix 16.2.

Supportive analysis will be run on the Per Protocol (PP) population and a sensitivity analysis will be conducted on the Full Analysis Set – as Randomized (FAS-AR).

The primary population for the assessment of safety will be the Safety population. Safety will be summarized and presented by randomization arm and according to product use pattern categories.

10.1 Full Analysis Set - As Randomized (FAS-AR)

The FAS-AR consists of all the randomized subjects with signed ICF who have both a valid baseline and at least one valid post-randomization value for one of the primary endpoints collected at V7, V10, or V16.

Subjects enrolled at sites that are terminated due to findings of non-compliance with good clinical practice (GCP) and/or with the protocol will be excluded from the FAS-AR.

Subjects included in the FAS-AR are analyzed as per randomized arm.

10.2 Full Analysis Set - As Exposed (FAS-EX)

The FAS-EX consists of all subjects in FAS-AR who have at least one record of reported product use diary post randomization.

The exposure assignment for the FAS-EX will be actual product exposure as explained in Section 4.1.2 “Definition of Product Use Pattern Categories”. In particular, the analysis on FAS-EX will be conducted using time-varying product use categories, where Month 1 – Month 6 will be analyzed using product use categories defined for]V4, V10[, and Month 7 – Month 12 data will be analyzed using the product use categories defined for]V4, V16[.

10.3 Per Protocol Population

The PP population is a subset of the FAS-EX and includes all randomized subjects who:

- Are subjects randomized to THS 2.2 who have a product use pattern category of THS-use or are subjects randomized to CC who have a product use pattern category of CC-use.
- Were included in the PP population in the original study.
- Do not make a quit attempt (see Section 6.3.3 "Adherence to Product Allocation").
- Have no major protocol deviation that impacts the overall subject evaluability (see Section 11.1"Major Protocol Deviations").
- Have data for at least one endpoint of the "smoker's health profile" which is not missing or excluded from the PP analysis at V7, V10, or V16, because of a major protocol deviation on product adherence and diary adherence (see Section 6.3.3 "Adherence to Product Allocation") impacting evaluability (See Section 11.1"Major Protocol Deviations").

Subjects included in the PP population are analyzed as per randomized arm.

10.4 Safety Population

The Safety population consists of all the subjects enrolled with signed ICF who have at least one valid value for a safety assessment. This analysis set excludes all subjects enrolled in sites terminated due to findings of non-compliance with GCP and/or with the protocol. In general, the exposure assignment for the Safety population will be by randomization arm and according to the product use pattern categories (]V4, V16[) as defined in Section 4.1.2 "Definition of Product Use Pattern Categories".

11 PROTOCOL DEVIATIONS

Protocol deviations are defined as deviations from any procedure defined in the Study Protocol, including but not limited to, as any violation of inclusion/exclusion criteria, mis-randomizations, assessments not performed according to protocol restrictions, or use of drugs that are known to affect the component of the "smokers' health profile".

Information following site monitoring and other reviews will be documented in the site visit reports, follow-up letters, audit documentation, or other manual review and subsequently recorded in an electronic data capture (EDC) system. Additional protocol deviations may be identified in the data review, these will also be recorded in the EDC system or imported into Study Data Tabulation Model (SDTM) datasets.

All deviations will be reviewed to determine their impact when subjects are assigned to analysis populations. Each deviation will be classified as major or minor; all major deviations will be further reviewed to determine whether or not the deviation impacts the evaluability of the results and therefore should result in the subject being excluded from the PP population.

11.1 Major Protocol Deviations

Subjects with major protocol deviations will be identified to determine whether they will be excluded from any of the analysis populations.

The categories for the major deviations may include, but are not limited to the deviations presented in Table 13.

Table 13: Definition of Major Protocol Deviations

Category	Description
Mis-randomization	Being administered the wrong product according to the randomization schedule or being randomized with an incorrect stratification variable.
Product misallocation	Subjects randomized to CC who had access to THS 2.2.
Product adherence	Non-adherence in the THS 2.2 arm during period]V4, V7],]V7, V10[or [V10, V16[is defined based on the reported product use as any of the following occurrence: - Use of more than 5 CC in a single day during the period. - Average THS 2.2 use is less than 70% of total THS 2.2 and CC use.
Entry Violation	Violation of inclusion/exclusion criteria.
Diary adherence	If less than 75% of the daily product use assessments over the period]V4, V7],]V7, V10[, [V10, V16[are available, or product use data is missing over a period of more than 7 consecutive days.
Duration of 24 hour collection	Total urine collection covering a period of less than 23 hours and more than 25 hours for assessments at Baseline, V7, V10, or V16, including significant missing collections within the 24 hour sample.

Among the above criteria, violations of inclusion criteria 1 or 2, or exclusion criteria 2, 16, 17, 19 or 21 of the original study or violation of the extension inclusion criterion 2 or exclusion criterion 4 will be considered as impacting the evaluability. Duration of 24h urine sample may be considered as not impacting the evaluability if collected at V7 or V10, or if collected over a period of at least 4 hours. The other criteria will be assessed for their impact on the PP population and evaluated during the pre-analysis data review meeting (Section 6.3.1 “Methods of Assigning Subjects to Product Arms”).

All major protocol deviations related to the original study and impacting evaluability will impact the overall subject evaluability leading to the exclusion from the PP population, apart

from those related to product adherence and diary adherence. Following major protocol deviation of product adherence or diary adherence impacting evaluability in period]V4,V7], all non-safety data at V7 will be flagged and excluded from the descriptive summaries and analysis for the PP population. The same approach will be used for descriptive summaries and analysis at V10 following a major protocol deviation of product adherence or diary adherence impacting evaluability in the]V7, V10[.

All protocol deviations occurring during the extension study and categorized as major and impacting evaluability will not directly impact the overall subject evaluability, however all non-safety data at Month 7 to Month 12 will be flagged and excluded from the descriptive summaries and analysis for the PP population.

11.2 Minor Protocol Deviations

The categories for the minor deviations may include, but are not limited to the deviations presented in Table 14. The assessment windows specified in the clinical study protocol are shown in Table 14.

Table 14: Definition of Minor Protocol Deviation Categories

Category	Description
Procedural violation	Violation of any study procedures not affecting safety or data evaluability.
Concomitant Medication	Use of any concomitant medications interfering with the “Smokers’ health profile” endpoints at Baseline, V7, V10, and V16. (See Appendix 2 of the study protocol)
Visit window deviation	See details in Table 15 for ambulatory visit window deviations. Visit window deviations for 24-hour urine are classified as major deviations.
Time missing	Assessment date or time is missing
Assessment missing	Assessment is missing
Visit missing	Scheduled visit not done

Table 15: Assessment Windows

Assessment	Nominal Time point(s)	Window
24-hour urine collection	V3, V7, V10, and V16	Start collection in the morning of the day prior to the visit, end after 24 ± 1 hours since the starting time
Ambulatory Visit	From V1 to V2	Not more than 42 days
	From V2 to V3	Not more than 5 days
	From V3 to V4	8 days ± 2 days
	From V4 to V5, and subsequent monthly visits	± 5 days with respect to the planned visit date.

12 PLANNED STATISTICAL METHODS

12.1 General Considerations

Data analysis will be performed using SAS® Version 9.2 or higher.

For analysis purposes, data collected in this extension study will be pooled with relevant data from the original study (ZRHR-ERS-09-US) as indicated in the following sections, with reference to Baseline and data collected during visits V1 to V10.

Data listings will be provided for all data collected as defined in the protocol, ordered by randomization arm, subject, Day, and study visit, unless otherwise stated. Summary statistics and statistical analysis will only be presented for data where detailed in this SAP. All unscheduled assessments will be included in the listings.

Markedly non-lognormally distributed BoExp and risk marker data will be transformed or analyzed by appropriate non-parametric methods with bootstrapping using the SAS procedure PROC SURVEYSELECT. The seed to be used in any analysis will be contained in the SAS output and will be different for each analysis.

This study has no formal pre-specified hypotheses associated with the study objectives. However, exploratory statistical hypothesis testing will be conducted to assess baseline compariablity (see Section 12.3 "Demographics and Other Baseline Characteristics") and to evaluate the THS 2.2 effect on selected endpoints (Section 12.7 "Planned Statistical Analyses"). All the THS 2.2 effect estimates will be accompanied by 95% CI.

12.1.1 Stratified Presentation

Selected data will be presented as summaries stratifed by sex (male and female) and average daily CC consumption level at V1 ("10-19 cig/day" vs. ">19 cig/day"), as described in this section.

Stratified presentations will be conducted by product use pattern categories within randomization arm for the FAS-EX at]V4, V16[for descriptive summaries of the following:

- Demographics (see 12.3 "Demographics and Other Baseline Characteristics") by sex.
- Product use (see 12.7.2.1 "Product Use") by sex and average daily CC consumption level at V1.
- "Smokers' health profile" endpoints (see 12.7.1.3.6 "Descriptive Analysis of Primary Endpoints") by sex and average daily CC consumption level at V1.

12.1.2 Subgroup Analyses

Subgroup analyses will be based on the final covariates that are included in the primary analysis model. Potential covariates for subgroup analyses are presented in Table 3.

12.1.3 Descriptive Statistics

In general, data are summarized by product use pattern category for FAS-EX and by randomization arm for the FAS-AR and PP populations for all visits during the 12-month randomized exposure period.

For the safety population, summaries will be in general produced by product use pattern category over the 12-month randomized exposure period, as described in Section 4.1.2.

For continuous data, summary statistics will include the number of subjects (n), the number and percent of subjects with missing data, the arithmetic mean, arithmetic standard deviation (SD), 95% confidence interval (CI) of the arithmetic mean, median, first and third quartiles, minimum, maximum; for log-normal data the geometric mean, geometric CV, and 95% CI of the geometric mean will be presented instead of arithmetic mean, SD, and 95% CI of the arithmetic mean. Post randomization summaries will include change from baseline apart from log-normal variables which will present % change and % relative change from baseline. For categorical data, frequency counts and percentages will be presented. For the calculation of summary statistics and statistical analysis, unrounded data will be used. Data listings will include all subject level data collected unless otherwise specified.

The following product labels will be used throughout the TFLs (Table 16):

Table 16: Product Labels

Product	Format used in TFLs	Order in TFLs
Tobacco Heating System 2.2	THS 2.2	1
Conventional cigarettes	CC	2

12.1.4 Handling of Dropouts or Missing Data (including outside the limits of quantification)

For the values of biomarkers and risk markers:

- Missing data at Baseline will be considered as missing completely at random, and only subjects with complete data for all the terms in the mixed model for repeated measures (MMRM) will be included in the analysis. Summary tables will include all available values.
- Missing data at Baseline for the propensity score analysis will be handled as described in Section 12.7.3.2 "Exposure/Response and Effect of Dual-Use on Smokers' Health Profile".
- Missing data after randomization will be considered as missing at random, and handled by means of MMRM approach.
- If more than 10% of the FAS-EX analysis population have missing data at Baseline or V16 for at least one of the primary endpoints, a sensitivity analysis will be conducted by means of a multiple missing imputation approach.

For laboratory parameters:

- Values below the lower limit of quantification (LLOQ) will be imputed using LLOQ/2. For values above the upper limit of quantification (ULOQ), the ULOQ will be used for calculation and reporting in summary tables.
- The number of values below LLOQ or above ULOQ will be presented in each summary table. If 50% or more data are below LLOQ or above ULOQ, only the number (%) of value below LLOQ or above ULOQ will be reported in the summaries, together with minimum (if no value below LLOQ is present) and maximum (if no value above ULOQ are present) of the observed values.
- For calculation of parameters derived as a ratio of two components (creatinine-adjusted 24-hour urine endpoints, Apo B/Apo A1) if the denominator is below LLOQ then the ratio will not be calculated and will be tabulated as "Not Calculated" in descriptive summaries and excluded from the main analysis. A sensitivity analysis

will be produced by including derived values using half LLOQ values for the denominator of the impacted endpoint for the FAS-EX.

- Missing data at Baseline will not be imputed.

For daily product use data:

- Missing product use data will be handled as specified in Section 16.3 "Data Handling and Derivation of Product Use Pattern Categories" and Section 12.7.1.3.3 "FAS-EX Supportive Analysis".

For calculation of questionnaires' scores:

- For MCEQ, domain or subscale scores will be derived by averaging the individual non-missing item scores if at least 50% are non-missing, otherwise they will be set to missing.

- For FTND questionnaire, the total score will be derived by summing the individual item scores if all items are non-missing; if any items are missing the total score will be set to missing.

- For SES questionnaire, observations with missing income or education will be excluded from the analysis of the composite SES.

- For assessment of cough, if no cough was reported at a given visit then the VAS score is imputed with zero for the calculation of change from baseline.

For Partial Dates:

- Partial dates will not be imputed for AEs, for medical history, and for concomitant medications, but assumptions will be made as follows to assign them to specific analysis categories:

Date information	AE Category	Disease Category	Medication Category
Missing or Partial date (*e.g.*, --May2012, or ----2011). If month/year is the same as, or later than the month and/or year of Screening.	Product-emergent	Concomitant disease	Concomitant medication
Partial date, (*e.g.*, --May2012, or ----2011). If month and/or year is earlier than the month and/or year of Screening.	Not product-emergent	Medical history	Prior medication

12.1.5 Insufficient Data for Analysis/Presentation

If there are no values or events at the general value then the break out should not be presented. For example if the number of AEs related to study procedure is zero then the presentation by severity of related events will not be produced.

Some of the TFLs will not have sufficient numbers of subjects or data for presentation. If this occurs, the blank TFL shell will be presented with a message printed in the center of the table, such as: "No applicable data for this summary."

For categories of continuous summaries that have <4 subjects no summaries will be shown.

12.1.6 Handling of Unplanned and Early Termination Assessment Data

Unscheduled assessments will be in general excluded from the analysis and summary statistics. However certain data planned for scheduled assessments are reported within unscheduled visit forms in the eCRF (e.g. repeated/delayed assessments). Given that it is not practical to re-assign them to the pertaining scheduled visits in the eCRF, these data will be accounted for in the analysis and summaries by applying the following re-mapping algorithm to each unscheduled visit:

- If Unscheduled visit date < enrollment date then analysis timepoint = Screening
- If Unscheduled visit date ≥ enrollment date and < randomization date then analysis timpoint = Enrollment
- If Unscheduled visit date ≥ randomization date then:
    1. Determine the unscheduled visit study day using the visit date until midnight.
    2. Determine the reference scheduled visit as the latest previous (≤) scheduled visit recorded and derive the nominal study day from Table 17.
    3. If unscheduled visit study day ≤ nominal study day of the reference scheduled visit + 30 days then unscheduled visit data will be assigned to the analysis timepoint of the reference scheduled visit (see Table 2); otherwise data will be remapped to the analysis time point using the time windows described in Table 17.

For example, someone missing lab assessments at V7 but performed them at study day 123 as recorded in unscheduled form, would have unscheduled lab assessments not mapped to Month 3 (because performed 31 days later than nominal V7 study day) but mapped to Month 4 (because the unscheduled study day is in the [107, 137] time window of Month 4).

Table 17: Mapping of Visits After Randomization

Analysis Time Point	Nominal Study Day (Study Visit and Week number)	Time Window (in Study Days) for Assessments
Day 1	1	= 1
Month 1	29 (V5, Week 4)	[2, 46]
Month 2	64 (V6, Week 9)	[47, 78]
Month 3	92 (V7, Week 13)	[79, 106]
Month 4	120 (V8, Week 17)	[107, 137]
Month 5	155 (V9, Week 22)	[138, 169]
Month 6	183 (V10, Week 26)	[170, 197]
Month 7	211 (V11, Week 30)	[198, 228]
Month 8	246 (V12, Week 35)	[229, 260]
Month 9	274 (V13, Week 39)	[261, 288]
Month 10	302 (V14, Week 43)	[289, 319]
Month 11	337 (V15, Week 48)	[320, 351]
Month 12	365 (V16, Week 52)	≥ 352

Unscheduled assessments will be labelled as unscheduled and presented together with the re-mapped analysis time point in the listings.

Early Termination Assessments will be mapped using the relative study day using the visit date until midnight. Given that Early Termination visits are in general conducted during scheduled visits, the mapping will be performed using time windows as defined in Table 17, apart from early termination visits before randomization (e.g. subjects enrolled and not randomized) which will be assigned to the Enrollment time point. Early termination visit assessments will be labelled as early termination visit and presented together with the re-mapped analysis time point in the listings.

Data from Unscheduled and Early Termination assessments will be combined with data from scheduled visits in the analysis and summary statistics only if there are not scheduled visit

data for the same parameters at that re-mapped analysis time point, and only if the data are scheduled to be collected at that analysis timepoint.

12.1.7 Multiple Comparisons / Multiplicity

This study has no formal pre-specified hypotheses associated with the study objectives. P-values and confidence intervals for all analyses are presented with no adjustment for multiplicity.

12.2 Disposition of Subjects

The number and percent of subjects will be summarized for the following categories: subjects screened, screen failures, enrolled, enrolled but not randomized, randomized, completed V7, completed V10, completed V16, and discontinued (Table 15.2.1.1).

Screen failures will be summarized by reason for screen failure. Unmet inclusion and exclusion criteria will be listed for each screened subjects having any unmet criterion, indicating the applied protocol version (Listing 15.3.1.2).

All subjects who discontinue the study (either main or extension) will be categorized by their primary reason for discontinuation and summarized by randomization arm for the enrolled population (Table 15.2.1.2). Disposition of subjects and reasons for discontinuation will also be summarized separately. Supportive listings will be provided (Listings 15.3.1.1 and 15.3.1.1.1).

The number and percent of subjects with protocol deviations, and the number of protocol deviations, will be summarized by randomization arm for the Randomized population (Listing 15.3.1.10; Table 15.2.1.3.1). Summaries will be broken down by main deviation category (major/minor), sub-categories and evaluability. Subjects will be counted once per deviation category, and can be counted for more than one deviation category.

The number and percent of randomized subjects included in each analysis population will be summarized by randomization arm for all Screened subjects (Table 15.2.1.3.2). For subjects excluded from the analysis populations this will be broken down by reason for exclusion. Subjects will be counted once per exclusion, and can be counted for more than one exclusion.

Supportive listings will be provided, including any additional comments for tests that are not performed to be included on the listings of individual data (Listings 15.3.1.8 and 15.3.2.4).

12.3 Demographics and Other Baseline Characteristics

The demographic variables and other baseline chracteristics such as age, sex, race, body weight, height, BMI and waist circumference will be summarized by 6-month and 12-month

product use pattern categories for the FAS-EX, and stratified as described in Section 12.1.1 for the 12-month product use pattern categories. In addition, the analysis will be presented for the FAS-EX for only those subjects who enrolled into the extension study for both 6-month and 12-month product use pattern categories. Analysis will also be presented by randomization arm for the Safety and PP populations (Listing 15.3.1.7; Tables 15.2.1.4.1.*, 15.2.1.4.2.*). The summary will be produced also for the FAS-AR if it is different from the Safety populaton post-randomization. Other baseline characteristics will also be performed as described below.

There will be no inferential analyses of baseline characteristic data performed; however hypothesis testing will be conducted to assess baseline comparability of THS-use and CC-use product use exposure groups at]V4, V10[and]V4, V16[(for all the covariates in Table 3) for the FAS-EX by means of Analysis of Variance method and chi-square tests for continuous and categorical variables, respectively (Tables 15.2.1.4.3.1 and 15.2.1.4.3.2). These results will be considered for the analysis as described in Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation".

Baseline comparability will also be conducted between subjects enrolled and not enrolled in the extension study among those who completed V10 (Table 15.2.1.4.4).

Subject responses, individual educational attainment, annual household income and composite classifications will be listed for the SES questionnaire (see Section 7.2.7 "Socio-Economic Status Questionnaire"). The number and percentage of subjects in each of the socio-economic status gradings will also be summarized (Listing 15.3.1.11; Tables 15.2.1.4.1.*, 15.2.1.4.2.*).

Subject response to items from the Prochaska questionnaire (see Section 7.2.4 "Prochaska 'Stage of Change' Questionnaire: Intention to Quit Smoking") will be provided in listings and summary tables (Listing 15.3.1.4; Tables 15.2.1.4.1.*, 15.2.1.4.2.*).

Smoking history as collected from the smoking history questionnaire (Section 7.2.3 "Smoking History Questions") will be listed (Listing 15.3.1.3) and summarized, as well as the number and percentage of subjects in the daily CC consumption level at V1 ("10-19 cig/day" vs. ">19 cig/day") (Tables 15.2.1.4.1.*, 15.2.1.4.2.*).

The number and percent of subjects in each category (THS 2.2/ CC/ No preference) of the answer to subjects' preference question (Section 7.2.6 "Product Preference Question") will be listed and summarized (Listing 15.3.1.5; Table 15.2.1.4.1.*, 15.2.1.4.2.*).

Other Baseline data on covariates (see Section 7.2.9 "Lifestyle Assessment") will be listed and summarized (Listing 15.3.1.12; Tables 15.2.1.4.1.*, 15.2.1.4.2.*).

12.4 Medical History and Concomitant Diseases

Medical history is defined as any condition that started and ended prior to V1. A concomitant disease is defined as any condition that is either detected at V1, and/or is still ongoing at V1.

Medical history and concomitant disease will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 18 or above and listed separately by randomization arm, System Organ Class (SOC) and Preferred Term (PT) within SOC.

Number and percent of subjects with Medical History and Concomitant diseases will be summarized by product use pattern categories (calculated over the entire 12-month ambulatory period), SOC and PT for the Safety population (Listing 15.3.1.9; Tables 15.2.1.5 and 15.2.1.6).

12.5 Measurements of Product Adherence

The product use electronic diary will serve as a tool to measure product adherence, as defined in Section 6.3.3 "Adherence to Product Allocation". The number of products used daily will be listed and described on the FAS-AR, PP, and Safety populations (Listing 15.3.2.1, 15.3.2.1.1, and 15.3.2.1.2; Tables 15.2.2.1.1-15.2.2.1.3), and the number and percentage of adherent subjects will be described on the FAS-EX, Safety, and FAS-AR populations overall and during periods]V4, V10[and [V10, V16[(Tables 15.2.5.1-15.2.5.3).

12.6 Other Data

Urine pregnancy results will be provided in Listing 15.3.1.6 and Attempt to quit smoking will be provided in Listing 15.3.1.13.

12.7 Planned Statistical Analyses

12.7.1 Primary Analyses

12.7.1.1 Hypotheses

No hypotheses are to be tested. The objective of the study is to determine the effect of THS 2.2 compared to CC at week 52 on the components of the "smokers' health profile" and to provide additional information to the results of the original study (ZRHR-ERS-09-US) for a prolonged exposure period.

12.7.1.2 Smokers' Health Profile Endpoints Evaluation

The following endpoints of the "smokers' health profile" values at Baseline, at V7, at V10, and at V16 will be log-transformed ($base_e$) prior to the analysis: sICAM-1, 11-DTX-B_2, 8-epi-PGF2α, COHb, and Total NNAL. HDL-C, WBC, and FEV_1 will be analyzed in the original scale. Concentration of urinary clinical risk endpoints and BoExp will be adjusted for creatinine, as defined in Section 7.1.1 "Clinical Risk Endpoints and Biomarkers of Exposure in Urine".

Clinical risk endpoints will be analyzed in the FAS-EX population for the subgroup comparing THS-use with CC-use categories, using a MMRM adjusting for sex, Caucasian origin, visit, value of the endpoint at Baseline and its interaction with visit, product use pattern category and its interaction with visit, and other Baseline covariates relevant for each specific clinical risk endpoint. Site will be included as a random effect [20]. Additional Baseline characteristics are defined for each endpoint of the "smokers' health profile" in Section 4.1.4 "Covariates for Smokers' Health Profile Analysis" and summarized in Table 3. Model will include terms for the "Defined Covariates" and for the subset of "Evaluated Covariates" selected if found to be significant at 10% level (between THS-use and CC-use) at Baseline for at least one of the 6-month (same as original study) or the 12-month product use categories (Table 15.2.3.1).

Modeling assumptions are evaluated and model fit is assessed by the analysis of residuals (model diagnostics) and by comparing the values predicted versus the observed endpoint values (calibration).

The SAS code for the mixed model to be used is shown below:

```
Proc mixed data=_data_ method=reml maxiter=200;
Class subject product_use site sex caucasian_origin visit;
Model endpoint = baseline sex caucasian_origin product_use visit
      product_use*visit baseline*visit <covariates> / ddfm=kr;
Random site;
Repeated visit / subject=subject type=un r rcorr;
Lsmeans product_use*visit / slice=visit alpha=0.05 diff cl pdiff;
Run;
```

The least squares (LS) means and estimate of the difference along with its 95% CI and 2-sided p-value will be presented in tables for HDL-C, WBC, and FEV_1 at V7, V10, and V16. For other clinical risk endpoints of the "smokers' health profile", results will be presented back transformed in the original scale as a ratio, as well as percent reduction relative to CC-use (i.e. 1 – THS-use:CC-use ratio) at V7, V10, and V16 (Table 15.2.3.1.1). THS-use vs CC-use will also displayed as forest plots at V7, V10, and V16 (Figure 15.1.1.1) for the FAS-EX population. In case of model convergence issues, this will be reported in the study report

and additional covariance structures will be investigated with the following order: heterogeneous compound symmetry (type=csh), heterogeneous toeplitz (type=toeph), heterogeneous autoregressive (1) (type=arh(1)), and variance components (type=vc).

As the smoking cessation effect will be estimated in the SA-SCR-01 study (NCT02432729), the analysis methods for the estimation of the preserved effect will be detailed in a separate analysis plan (cross-study analysis).

12.7.1.3 Supportive Analysis

12.7.1.3.1 PP Analysis (Subjects Adherent to the Assigned Product)

The primary analysis of the "smokers' health profile" endpoints (see Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation") will be repeated as a secondary analysis for the PP population (Table 15.2.3.2.2) using a 95% CI and 2-sided p-value. Subjects will be analyzed as randomized by means of a MMRM model including the same adjustment covariates used for the primary analysis (see Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation"), except that randomization arm will be used instead of product use category.

Forest plots of THS 2.2 versus CC arms will be displayed at V7, V10, and V16 for the PP population (Figure 15.1.1.2).

12.7.1.3.2 FAS-AR Sensitivity Analysis

A sensitivity analysis will be conducted on the "smokers' health profile" endpoints for the FAS-AR (Table 15.2.3.2.3) using a 95% CI and 2-sided p-value to compare the THS 2.2 and CC arms. Subjects will be analyzed as randomized by means of a MMRM model including the same adjustment covariates used for the primary analysis (see Section 12.7.1.2 Smokers' Health Profile Endpoints Evaluation"), except that randomization arm will be used instead of product use category. The FAS-AR results will be interpretated taking into account those observed from the sensitivity analysis described in Section 12.7.3.2 "Exposure/Response and Effect of Dual-Use on Smokers' Health Profile".

Forest plots of THS 2.2 versus CC arms will be displayed at V7, V10, and V16 for the FAS-AR (Figure 15.1.1.3).

12.7.1.3.3 FAS-EX Supportive Analysis

A supportive analysis will be conducted on the "smokers' health profile" endpoints for the FAS-EX (Table 15.2.3.2.1) to compare the Dual-use versus CC-use at V7, V10, and V16. Subjects will be analyzed by means of a MMRM model including the same adjustment covariates used for the primary analysis (see Section 12.7.1.2 Smokers' Health Profile

Endpoints Evaluation"). The estimates of the difference/percent reduction (and 2-sided p-values) will be presented together with 95% CI for Dual-use vs CC-use and also displayed as forest plots at V7, V10, and V16 (Figure 15.1.1.1) for the FAS-EX population.

12.7.1.3.4 Missing Product Use Sensitivity Analysis

A Missing product use sensitivity analysis will be produced for the FAS-EX population using a 95% CI and 2-sided p-value to compare THS-use versus CC-use, by re-analyzing the primary endpoints re-classifing all subjects based on the available data in the diary (Table 15.2.3.1.2) with no imputation (see Section 16.3.2).

12.7.1.3.5 Sensitivity Analysis for Subjects Enrolled into Extension Study

A sensitivity analysis will be produced for the subset of FAS-EX subjects who enrolled into the extension study using a 95% CI and 2-sided p-value to compare THS-use versus CC-use and Dual-use vs CC-use (Table 15.2.3.1.3).

12.7.1.3.6 Multiple Imputation Sensitivity Analysis

A multiple imputation sensitivity analysis will be conducted on the "smokers' health profile" endpoints for the FAS-EX (Table 15.2.3.1.4) using a 95% CI and 2-sided p-value to compare THS-use versus CC-use at Months 6 and 12. The analysis will only be performed for endpoints which have more than 10% missing data at baseline or Month 12 for those in the FAS-EX population.

First, missing data will be imputed separately for each product use category (]V4, V16[) using the Markov Chain Monte Carlo multiple imputation method with terms including baseline value, Month 3 value, Month 6 value, Month 12 value, sex, Caucasian origin, and the same adjustment covariates for each specific clinical risk endpoint as included in the primary analysis. For log-transformed parameters, log-transformed values will be used during imputation.

Then, for each imputed dataset, adjusted LS means and CIs will be determined from a mixed effects model conducted on original Month 6 or Month 12 values with baseline value, sex, Caucasian origin, product use pattern category (Month 6 or Month 12), and other baseline covariates relevant for each specific clinical risk endpoint as fixed effect factors and site as a random effect. For log-transformed parameters, log-transformed values will be used in the model, and adjusted geometric LS means and CIs will be determined.

Finally, the results across imputations are combined to produce final inferential estimates. In case of model convergence issues, this will be reported in the study report and additional covariance structures will be investigated with the following order: heterogeneous compound symmetry (type=csh), heterogeneous toeplitz (type=toeph), heterogeneous autoregressive (1) (type=arh(1)), and variance components (type=vc).

12.7.1.3.7 Sensitivity Analysis for Derived Endpoints Including LLOQ Values in Denominator

A sensitivity analysis will be produced for the 24-hour urine creatinine-adjusted "smokers' health profile" endpoints in which derived values will be calculated using half LLOQ creatinine values for the denominator in cases where the creatinine value is below LLOQ. This sensitivity analysis will be performed for the FAS-EX population using a 95% CI and 2-sided p-value to compare THS-use versus CC-use and Dual-use vs CC-use (Table 15.2.3.1.5).

12.7.1.3.8 Descriptive Analysis of Primary Endpoints

The primary endpoints of the "smokers' health profile" (HDL-C, WBC, sICAM-1, 11-DTXB2, 8-epi-PGF2α, COHb, FEV_1, Total NNAL) will be summarized as detailed in Section 12.1.3 "Descriptive Statistics" for the FAS-EX, FAS-AR, and PP Populations (Tables 15.2.4.1.*). The primary endpoints will also be summarized by sex and average daily CC consumption at V1 for the FAS-EX (Tables 15.2.4.1.1.*).

The listing of the urinary clinical risk endpoints and BoExp will include the concentration adjusted for creatinine and the percent change from Baseline in the concentration adjusted for creatinine (Listings 15.3.3.1-15.3.3.3).

The listing and summary of the COHb data will include the concentration, the percent change from Baseline in the levels, and a flag for whether a subject's COHb was <2%. The listing of CO will flag subjects with CO level > 10 ppm.

12.7.2 Secondary Analyses

12.7.2.1 Product Use

The number of CC or THS Tobacco Sticks used daily (as reported on the self-reported product use electronic diary) will be described on the FAS-EX (for both 6-month and 12-month product use categories) and PP population and on the Safety Population. Summaries will include product use of other nicotine/tobacco-containing products and overall tobacco product grouping CC, THS Tobacco sticks, and other tobacco product (Listings 15.3.2.1-15.3.2.3; Tables 15.2.2.1.1-15.2.2.1.3 and 15.2.2.1.1.1). Summaries will also be presented by

sex and average daily CC consumption at V1 for the 12-month product use categories for the FAS-EX (Tables 15.2.2.1.1.2 and 15.2.2.1.1.3). In addition, summaries will be presented for the 6-month product use categories for the FAS-EX (15.2.2.1.1.4).

Product use data will be summarized by the monthly exposure period by study arm on the Safety Population (Table 15.2.2.1.3.1).

In addition, the number and percent of subjects in each product use pattern category will be summarized in 6-month intervals (]V4, V10[, [V10, V16[) and in 12-month interval (]V4, V16[) for the FAS-EX, Safety, and FAS-AR populations (Tables 15.2.5.1-15.2.5.3). Summaries will also be presented by sex and average daily CC consumption at V1 for the FAS-EX (Tables 15.2.5.1.1 and 15.2.5.1.2). In addition, a summary of detailed product use patterns will be produced for the FAS-EX (Table 15.2.5.4).

The distribution of product use will also be summarized for the THS-use, Dual-use, and CC-use by cross-tabulation of the frequency (number and percent) of subjects for the categories of the following two variables for the FAS-EX population (Table 15.2.4.15):

- Average number of daily CC smoked over the 12 month period.
- Average number of daily THS Tobacco Sticks used over the 12 month period.

Categories will be defined as “<1”, “1-4”, “5-10”, “11-20”, “21+”.

12.7.2.2 Reduction of Exposure to Biomarker of Exposure

The level of the BoExp (CO, Total NNN) and their percent change from Baseline will be listed and summarized at V7, V10, and V16 (Listings 15.3.3.1, 15.3.3.2, and 15.3.3.3.1; Tables 15.2.4.2.1-15.2.4.2.3). All BoExp parameters will be log-transformed except CO.

The reduced exposure will be demonstrated if a statistically significant reduction of biomarkers level in THS-use compared to CC-use are observed. To this aim, the MMRM analysis approach adopted in Section 12.7.1.2 “Smokers’ Health Profile Endpoints Evaluation” will be used on log-transformed Total NNN to compare the geometrical means ratio between THS-use and CC-use at V7, V10, and V16, using a 5% Type I error with no adjustment for multiplicity (Tables 15.2.3.3.1-15.2.3.3.3). Exhaled CO will be analyzed on the original scale and test the LS mean difference between THS-use and CC-use at V7, V10, and V16. The models will include the same adjustment covariates used for the primary analysis of COHb and Total NNAL. Site will be included as a random effect. The ratios/differences will also be compared for Dual-use versus CC-use, using a 5% Type I error. The estimates of the difference/percent reduction will be presented together with 95% CI for THS 2.2 versus CC arms for the FAS-AR and PP populations.

All summaries and analysis will be produced for the FAS-EX, FAS-AR, and PP populations.

In addition, a sensitivity analysis will be produced for the Total NNN creatinine-adjusted endpoint in which the derived values will be calculated using half LLOQ creatinine values for the denominator in cases where the creatinine value is below LLOQ. This sensitivity analysis will be performed for the FAS-EX population to compare THS-use versus CC-use and Dual-use vs CC-use (Table 15.2.3.3.1.1).

12.7.2.3 Nicotine Exposure

The level of Neq in 24-hour urine adjusted for creatinine, Nicotine and Cotinine in plasma will be listed and summarized over time (Listings 15.3.3.4 and 15.3.3.4; Table 15.2.4.3.*).

The MMRM analysis approach adopted in Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation" will be used on log-transformed Neq in 24-hour urine adjusted for creatinine, Nicotine, and Cotinine levels to compare the geometrical means ratio between THS-use and CC-use at V7, V10, and V16, using a 5% Type I error with no adjustment for multiplicity. The models will include the same adjustment covariates used for the primary analysis of COHb and Total NNAL. Site will be included as a random effect. The geometrical mean ratios will also be tested for Dual-use versus CC-use, using a 5% Type I error. The estimates of the percent reduction will be presented together with 95% CI for THS 2.2 versus CC arms. All summaries and analysis will be produced for the FAS-EX, FAS-AR, and PP population (Tables 15.2.4.4.*).

In addition, a sensitivity analysis will be produced for the Neq creatinine-adjusted endpoint in which the derived values will be calculated using half LLOQ creatinine values for the denominator in cases where the creatinine value is below LLOQ. This sensitivity analysis will be performed for the FAS-EX population to compare THS-use versus CC-use and Dual-use vs CC-use (Table 15.2.4.4.1.1).

12.7.2.4 Change in biological and functional markers

12.7.2.4.1 Lung Function and Cardiovascular diseases markers

Lung function and Cardiovascular data will be collected at V3, V7, V10, and V16.

Lung function data includes spirometry (FEV_1, FVC, FEV_1/FVC, FEF 25-75, and bronchodilator reversibility in FEV_1) and lung volume data (FRC, VC, TLC, IC and RV), and the brand (trade) name and dose of the bronchodilator. Spirometry data will also be collected at V1. At V1, V10, and V16, pre and post-bronchodilator spirometry test will be conducted, while at V3 and V7 only post-bronchodilator test will be performed. Spirometry predicted values will be standardized to the National Health and Nutrition Examination

Survey III predicted set [21], using the formula by Hankinson et al. [22]. Predicted values for Lung function parameters will be derived by a central provider.

Cardiovascular data include the following:

- MPO, Apo A1 and Apo B, Apo B/Apo A1, LDL-C, hs-CRP in serum.
- Fibrinogen, and homocysteine in plasma.
- Platelet count, and HbA1c in whole blood.
- Albumin in urine (expressed as concentration adjusted to creatinine and as quantity excreted in 24h).
- Blood Pressure, weight, waist circumference.

The level of biological and functional markers and their change from Baseline will be listed, and summarized at V7, V10, and V16 (Listings 15.3.3.1-15.3.3.3, 15.3.6.9, and 15.3.6.11; Tables 15.2.4.5.1-15.2.4.5.3, and 15.2.4.7.1-15.2.4.7.3). RV will be summarized descriptively only. FEV_1 %pred (post-bronchodilator) will not be analyzed or presented descriptively with the other lung function data since it is already summarized as a primary endpoint. Albumin in urine expressed as quantity excreted in 24h will be listed only.

The MMRM analysis approach adopted in Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation" will be used to compare between THS-use and CC-use at V7, V10, and V16, using a 5% Type I error with no adjustment for multiplicity. Lung function data, apart from RV and FEV_1 %pred (post-bronchodilator), will be analyzed in the original scale (Tables 15.2.4.6.1-15.2.4.6.3 and 15.2.4.8.1-15.2.4.8.3), and cardiovascular data will be analyzed in orginal scale for Apo A1, Apo B, Apo B/Apo A1, LDL-C, HbA1c, blood pressure, weight, and waist circumference, and in logarithmic scale for albumin in urine (expressed as concentration adjusted to creatinine), MPO, hs-CRP, fibrinogen, homocysteine, and platelet count.

The models will only include terms for visit, sex, Caucasian origin, age, smoking intensity, baseline level, product use pattern category, and the interactions of visit with both baseline level and product use category. Site will be included as a random effect. The ratios/differences will also be compared for THS-use versus CC-use and Dual-use versus CC-use, using a 5% Type I error. The estimates of the difference or ratio/percent reduction will be presented together with 95% CI. These estimates will be produced for THS 2.2 versus CC arms for the FAS-AR and PP populations.

All summaries and analyses will be performed on the FAS-EX, FAS-AR, and PP populations.

In addition, a sensitivity analysis will be produced for the albumin in urine creatinine-adjusted endpoint and the Apo B/Apo A1 endpoint in which the derived values will be calculated using half LLOQ values for the denominator in cases where the value is below LLOQ. This sensitivity analysis will be performed for the FAS-EX population to compare THS-use versus CC-use and Dual-use vs CC-use (Table 15.2.4.8.1.1).

12.7.2.5 Cough Symptoms

Cough symptoms will be assessed by means of the cough questionnaire, as described in Section 7.2.8 "Cough Questionnaire".

Descriptive analysis for change from Baseline will be presented for the VAS score and evaluating the level of cough bother at V7, V10, and V16. For the calculation of the change from Baseline, missing VAS values will be imputed with zero when no cough was reported.

The responses to the individual items, including the VAS score, change from Baseline, 3 Likert scales measuring the intensity and frequency of cough, and the amount of sputum production will be listed and summarized using the number of subjects who reported to experience cough as denominator (Listings 15.3.3.5; Tables 15.2.4.9.1 and 15.2.4.9.2).

The number and percentage of subjects reporting a cough will be summarized by randomization arm and product use pattern categories.

The answers to the open question related to any other important observation will be listed.

In addition, a logistic mixed-effects regression model for repeated measures will be used to compare the percentages of subjects reporting the need to cough between THS-use and CC-use at V7, V10, and V16 (Table 15.2.4.10). The model will include terms for visit, sex, Caucasian origin, age, smoking intensity, baseline cough, product use pattern category, and the interactions of visit with both baseline cough and product use category. Site will be included as a random effect. The point and 95% CI of LS means of the proportion of subjects reporting need to cough and their odds ratio will be presented for THS-use versus CC-use and also for Dual-use versus CC-use.

The SAS code for the logistic regression model to be used is shown below:

```
Proc glimmix data=_data_ method=rspl maxopt=200;
Class subject baseline product_use site sex caucasian_origin
visit;
Model endpoint = baseline sex caucasian_origin age
      smoking_intensity product_use visit baseline*visit
      product_use*visit / dist=binary link=logit ddfm=kr;
```

```
Random site;
Random visit / subject=subject type=un residual;
Lsmeans product_use*visit / slice=visit alpha=0.05 diff cl
pdiff ilink or;
Run;
```

In case of model convergence issues, this will be reported in the study report and additional covariance structures will be investigated with the following order: heterogeneous compound symmetry (type=csh), heterogeneous toeplitz (type=toeph), heterogeneous autoregressive (1) (type=arh(1)), and variance components (type=vc). All descriptive summaries will be performed on the FAS-EX and PP populations and the logistic regression will be performed on the FAS-EX.

12.7.2.6 Subjective Effects of Smoking

12.7.2.6.1 Modified Cigarette Evaluation Questionnaire

The MCEQ assessment is described in Section 7.2.2 "Modified Cigarette Evaluation Questionnaire".

Descriptive analysis for change from Baseline will be presented for the five domain scores at V7, V10, and V16. Scores and changes from Baseline will be listed and summarized (Listing 15.3.4.1; Tables 15.2.4.11.1 and 15.2.4.11.2). The answers to the individual questions will be included in the listings.

In addition, the MMRM analysis approach adopted in Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation" will be used for the five domain scores to evaluate the mean difference between THS-use and CC-use at V7, V10, and V16, using a 5% Type I error with no adjustment for multiplicity (Tables 15.2.4.12.1 and 15.2.4.12.2). The analysis will also be conducted on THS 2.2 vs CC arm. The models will only include terms for visit, sex, Caucasian origin, age, smoking intensity, baseline score, product use pattern category, and the interactions of visit with both baseline score and product use category. Site will be included as a random effect. The mean differences will also be described for Dual-use versus CC-use categories, using a 5% Type I error. The estimates of the difference will be presented together with 95% CI for THS 2.2 versus CC arms.

All summaries and analyses will be performed on FAS-EX and PP populations.

12.7.3 Exploratory Analysis

12.7.3.1 Questionnaires

12.7.3.1.1 FTND Questionnaire

The FTND (Section 7.2.1 "Fagerström Test for Nicotine Dependence (FTND)") score, the number and percent of subjects in each category of the FTND score (Mild/ Moderate/ Severe) and the response to the first item of the FTND ("Within 5 minutes", "6 to 30 minutes", "31 to 60 minutes", "After 60 minutes") will be summarized and listed at V1, V10 and V16 (Listing 15.3.5.1; Tables 15.2.4.13.1 and 15.2.4.13.2).

The change from V1 in the FTND score category and response to the first item of the FTND will be presented in a shift table.

All summaries and shift tables will be performed on the FAS-EX (using the product use pattern category) and PP population.

12.7.3.1.2 Intent to Use THS 2.2 Questionnaire

Baseline for the Intent to Use Questionnaire is the assessment collected at Visit 4 since this is the only pre-randomization collection. The shift from Baseline will be calculated for each item of the ITUQ (Section 7.2.5 "Intent to Use THS 2.2 Questionnaire") at V10 and V16. Item scores and shift from Baseline will be summarized by product use category and listed (Listing 15.3.5.2; Tables 15.2.4.14.1 and 15.2.4.14.2). For items 1-6, the shifts will be presented using the categories of 'Positive intention' and 'Not positive intention' as defined in Section 7.2.5 "Intent to Use THS 2.2 Questionnaire".

All summaries will be performed on the FAS-EX and PP populations.

12.7.3.2 Exposure/Response and Effect of Dual-Use on Smokers' Health Profile

To describe the effect of dual-use on the components of the "smokers' health profile" the exposure-response relationship between the CC, THS 2.2 exposure and the "smokers' health profile" endpoints level will be evaluated by means of a propensity score approach, based on the method proposed by Follmann [23].

This approach aims to estimate the effect of THS 2.2 among those subjects who would adhere to the THS 2.2, using an ITT-type approach. It is based on the assumption that the subject's propensity to adhere to THS 2.2 could be estimated based on baseline covariates, and that randomization ensures that the overall population of subjects who would adhere to

THS 2.2 will be about the same in the two arms, even though this proportion is unobservable in the CC arm. Also by randomization, any relationship between product adherence and baseline covariates observed in the THS arm would be about the same in the CC arm, had the latter group been randomized to the THS 2.2 product.

The analysis will be performed in a two-step process. First a propensity score (PS) model is developed on the FAS-AR using baseline covariates to predict adherence to THS 2.2 as observed in the THS-use pattern category (]V4, V16[) in the THS 2.2 arm, by means of a logistic regression model.

```
proc logistic data = _THSarmData_;
model product_use(event='THS-use') = <model terms> /
stepwise=selection slentry=0.157 slstay=0.157;
run;
```

The following variables will be considered for their inclusion in the model by means of the stepwise selection method with entry and removal alpha levels of 15.7%, which approximately corresponds to the use of the Akaike Information Criterion [24]:

- Demographics and baseline characteristics (see Table 15.2.1.4.3).
- ITUQ items (items 1 to 6 will use the mapped categories of 'Positive intention' and 'Not positive intention' while item 7 will use the original 6-point scale).
- SES sub-scores (educational attainment and annual household income).
- Baseline of all primary endpoints and Neq; for sICAM-1, 11-DTX-B_2, 8-epi-PGF2α, COHb, Total NNAL, and Neq, the value will be natural log-transformed while HDL-C, WBC, and FEV_1 will be analyzed in the original scale.

In case of missing baseline data, the median of non-missing baseline values across all FAS-AR subjects in the THS 2.2 arm will be imputed for each continuous variable and the mode of non-missing baseline values across all FAS-AR subjects in the THS 2.2 arm will be imputed for each categorical variable.

The final model will be fit using the following SAS code. Goodness of fit diagnostics and discrimination ability of the fitted model based on the area of the receiver operating curve will be provided in statistical output listings.

```
proc logistic data = _THSarmData_ outmodel=_psmodel_;
model product_use(event='THS-use') = <model terms> /
clodds=wald lackfit outroc = ps_r;
```

```
output out= ps_p XBETA=ps_xb STDXBETA= ps_sdxb PREDICTED =
ps_pred;
run;
```

The variables included in the final PS model together with the odds ratios and 95% CIs will be tabulated (Table 15.2.4.16). The fitted model is used to predict scores on both the THS 2.2 and CC arms, using the following SAS code:

```
proc logistic inmodel=_psmodel_;
  score  data = _alldata_ out=_predData_;
run;
```

In case of missing values in covariates included in the final PS model, the same imputation approach used for PS model fitting will be adopted. For the scoring of subjects in the CC arm, the median and mode values used for imputation will be derived only from subjects in the CC arm of FAS-AR. Histograms of the predicted PS scores will be presented by randomization arm (Figure 15.1.2.1) and individual PS scores will be reported in the statistical output listings.

The second step consists of a MMRM model analysis conducted on the FAS-AR with the model terms noted below (including the same adjustment covariates used for the primary analysis - see Section 12.7.1.2 "Smokers' Health Profile Endpoints Evaluation") and including the interaction of predicted PS values fit as a natural cubic spline with the randomization arm and the interaction of predicted PS values fit as a natural cubic spline with visit (Table 15.2.4.17). The THS 2.2 vs CC effect at V10 and V16 will be tabulated for PS levels equal to 100%, 75%, 50%, and 25% and displayed as a forest plot (Figure 15.1.2.2). A PS*randomization arm interaction p-value (via the spline interaction) of 15.7% or less will be considered as indicative of a statistically significant THS effect based on the level of adherence to THS.

The SAS code for the mixed model to be used is shown below:

```
Proc glimmix data=_data_ method=rspl maxopt=200;
Class subject rand_arm site sex caucasian_origin visit;
Effect spl = spline(P_THS-use / naturalcubic);
Model endpoint = baseline sex caucasian_origin rand_arm spl
      visit    baseline*visit    rand_arm*visit    visit*spl
      rand_arm*spl <covariates> / ddfm=kr;
Nloptions technique=nrridg;
Random site;
random visit / subject=subject type=un residual;
Lsmeans  rand_arm*visit  /  at  P_THS-use=xxx  slice=visit
alpha=0.05 diff cl pdiff;
Run;
```

where P_THS-use=xxx is set to the 1, 0.75, 0.5, 0.25 for each PS level. Note that 'method=rspl' is equivalent to using 'method=reml' with proc mixed. Also, 'nloptions technique=nrridg' solves the iterative optimization problem by means of a ridge-stabilized Newton-Raphson algorithm as used by proc mixed. In case of model convergence issues, this will be reported in the study report and additional covariance structures will be investigated with the following order: heterogeneous compound symmetry (type=csh), heterogeneous toeplitz (type=toeph), heterogeneous autoregressive (1) (type=arh(1)), and variance components (type=vc).

12.7.4 Safety Evaluation

Safety variables monitored in this study include: AEs; spirometry (clinical interpretation and COPD categories) and respiratory symptoms (cough assessment VAS and Likert scales); vital signs (systolic and diastolic blood pressure, pulse rate, respiratory rate); ECG data; concomitant medications, clinical chemistry, hematology, urine analysis safety panel; physical examination, body weight, and BMI.

12.7.4.1 Safety Reporting

All safety data collected during the study will be provided in listings by randomization arm, site, subject, and product use pattern category defined over the 12-month period. Safety study periods are defined as follows:

- Screening: [Screening to V3[
- Product Trial: [V3 to V4]
- Exposure:]V4 to V16]
- Safety Follow-up: [Start to End of Safety Follow-up], where
    - the Start of Safety Follow-up is after Completion of V10 for subjects who complete the main study but do not enroll in the extension study, after Completion of V16 for completers of the extension study, or after the latter of the discontinuation date and the Early Termination visit date for non-completers

All summaries for safety parameters will be conducted on the Safety Population. Unless otherwise specified, summaries will be produced overall by randomization arm and product use pattern categories (see Section 4.1.2 "Definition of Product Use Pattern Categories").

Safety data collected during the Product Trial period will be summarized by randomization arm in a separate table, including subjects in the Safety population who are not randomized.

AEs occurring during the safety follow-up will be summarized accounting for the number of subjects entered in the safety follow-up period.

Unless otherwise specified, safety data collected during the Screening period will be provided in listings only.

12.7.4.2 Adverse Events

A product emergent AE is defined as an AE that occurs at Enrollment or later. AEs that occurred at Enrollment will be considered product emergent only if flagged as having started after first THS 2.2 use in the eCRF. All other AEs will not be summarized but provided in listings only. Additionally, product emergent AEs flagged as having occurred after THS 2.2 use or with a start date greater than any recorded THS 2.2 use date will be summarized during the Product Trial period.

Any new, clinically relevant, abnormal finding or worsening of a pre-existing condition/concomitant disease detected during the study including the safety follow-up period will be documented as an AE and/or SAE as described in the Safety Management Plan.

All AEs occurring from the signing of informed consent will be recorded electronically. AEs collected before product use will be provided in listings only.

12.7.4.2.1 All Adverse Events

General AEs summary tables will be presented for:

- The number of events and the number and percentage of subjects reporting at least one AE.
- The number of events and the number and percentage of subjects reporting at least one study product–related AE, broken down by product relatedness (related to THS 2.2/ CC) and expectedness (expected for THS 2.2 / CC).
- The number of events and the number and percentage of subjects with at least one AE leading to study discontinuation.
- The number of events and the number and percentage of subjects reporting at least one AE broken down by severity including each subject only once with the worst severity.
- The number of events and the number and percentage of subjects reporting at least one SAE.
- The number of events and the number and percentage of subjects reporting at least one AE leading to any action taken, broken down by action taken related to the product (product use interrupted, product use reduced, product use stopped, not applicable, none), treatment given (yes, no), study discontinuation, and other action taken.
- The number of events and the number and percentage of subjects reporting at least one AE related to study procedure.

An additional summary table of AEs will be also presented with a breakdown of the number of events, as well as the number and percentage of subjects reporting each AE, categorized by SOC and PT coded according to the MedDRA (version 18.0 or latest). If a subject has more than one occurrence of the same AE, the subject will be counted only once within a PT with the worst occurrence based on the presentation (*e.g.*, for presentation by severity = most severe, for presentation by relationship = most related). The number of events will be counted within each applicable category without regard to worst occurrence. For example, a subject who has the same AE three times will be counted three times in the number of events for that AE. Missing information on the intensity of AE will be counted as severe when determining the subject counts. Missing information on the relationship will be counted as related when determining the subject counts. Missing intensity and relationship will display as missing in the summary outputs when determining the number of events.

AE tables will be summarized by randomization arm and product use pattern category for the overall 12-month interval, and by randomization arm by monthly visits (Table 15.2.6.3.1). All AEs will be listed by randomization arm and product use pattern category (Listing 15.3.6.1; Tables 15.2.6.1- 15.2.6.12).

12.7.4.2.2 Serious Adverse Events (Including Deaths)

Summary tables of SAEs will be presented using the same approach as for AEs by randomization arm (see Section "All Adverse Events"), and including the number of events and the number and percentage of subjects reporting at least one SAE broken down by seriousness criteria (fatal, life-threatening, requires hospitalization, results in disability/incapacity, congenital anomaly/birth defect) (Table 15.2.6.13). SAEs will also be summarized by product use pattern category, if there are at least 10 SAEs (Table 15.2.6.14).

SAEs will also be listed in separate listings by randomization arm and product use pattern category (Listing 15.3.6.2).

12.7.4.2.3 Adverse Events Leading to Discontinuation

Summary tables of AEs leading to discontinuation will be presented by randomization arm, using the same approach as for AEs (see Section "All Adverse Events"). AEs leading to discontinuation will also be summarized by product use pattern categories if there are at least 10 AEs leading to discontinuation (Tables 15.2.6.15 and 15.2.6.16).

AEs leading to discontinuation will also be listed in separate listings ordered by subject, randomization arm, and product use pattern category (Listing 15.3.6.3).

12.7.4.2.4 Laboratory Abnormalities

The shift in toxicity grades from Baseline to worst grade recorded during the 12-month randomized exposure period will be presented in tables for the clinical chemistry, hematology and urinalysis parameters (Tables 15.2.6.18-15.2.6.23). Details related to the toxicity grading of laboratory abnormalities are available in Section 12.7.4.3 "Clinical Laboratory Evaluation".

12.7.4.2.5 THS 2.2 Device Events

All events relating to the device type will be listed, including event description, device type the event relates to, severity of event, AE relationship, proposed solution and onset/stop dates/times (Listing 15.3.6.5). Device events will be classified according to C54451/Medical_Device_Problem_Codes_FDA_CDRH [25].

A summary table (Table 15.2.6.17) of device events will be presented by randomization arm for both the run-in period and 12-month randomized exposure period, including:

- Number of device events and the number and percentage of subjects reporting at least one device event.
- Number of device events and the number and percentage of subjects categorized by severity of device event (minor, major).
- Number of device events and the number and percentage of subjects categorized by AE relationship (related, not related).
- Number of device events and the number and percentage of subjects categorized by event description.

All device events will be listed; data collected during Screening will not be summarized.

12.7.4.3 Clinical Laboratory Evaluation

Hematology, clinical chemistry and urine analysis parameters to be assessed at V1, V3, V7, V10, and V16 are listed in Table 18.

Table 18: List of Laboratory Safety Parameters

Hematology	Clinical Chemistry	Urine analysis
Hematocrit	Albumin	pH
Hemoglobin	Total protein	Bilirubin
Mean corpuscular hemoglobin (MCH)	Alkaline phosphatase (AP)	Glucose
	Alanine aminotransferase	Nitrite

Table 18: List of Laboratory Safety Parameters

Hematology	Clinical Chemistry	Urine analysis
Mean corpuscular hemoglobin concentration(MCHC) Mean corpuscular volume(MCV) Platelet count Red blood cell (RBC) count* White blood cell (WBC) count* Differential WBC count: • Neutrophils • Basophils • Eosinophils • Lymphocytes • Monocytes	(ALT) Aspartate aminotransferase(AST) Blood urea nitrogen (BUN) Creatinine Gamma-glutamyl transferase GGT) Fasting Glucose Lactate dehydrogenase(LDH) Potassium Sodium Total bilirubin Direct bilirubin Indirect bilirubin** Total cholesterol(TC) Triglycerides(TG)	Red blood cell traces Protein Specific gravity

*RBC and WBC morphology tests might be performed in case of abnormal values.
**Indirect bilirubin calculated as Total bilirubin – Direct bilirubin

Any clinical safety laboratory test result that is outside of the normal reference range will be reviewed by the Principal Investigator or designee and assessed for its clinical relevance. If the abnormal result is considered to be of clinical relevance, then it must be recorded as a concomitant disease at V1, or if not present at V1, as an AE, or still linked to an AE or to a concomitant disease, during the study. If the condition worsens from screening to after product exposure it will be recorded as an AE or linked to an AE.

The grading scheme used in the Common Terminology Criteria for Adverse Events and Common Toxicity Criteria ([CTCAE] version 4.03) will be used by the Principal Investigator or designee to assess abnormal laboratory values. These CTCAE grades will be derived programmatically in the creation of the datasets.

Laboratory data will be summarized by randomization arm and listed at Screening, Baseline, and V5 to V16 together with changes from Baseline (Listings 15.3.6.6-15.3.6.8; Tables 15.2.6.18-15.2.6.23). The number and percentage of subjects with normal results, high/low results (with respect to the reference range) and abnormal clinically significant results (as defined by Principal Investigator's or designee's comments) will be tabulated for laboratory parameters by randomization arm and product use pattern category.

Listings for the clinical laboratory will be provided by randomization arm. Data will include the following information: normal/high/low, abnormal clinically significant, Principal Investigator's or designee's comments, change from Baseline, and CTCAE grade. Only CTCAE grades greater than zero will be presented.

Shift tables based on the change from Baseline in CTCAE grades to the worst CTCAE grade after Baseline will also be produced by laboratory parameter, randomization arm and product use pattern category.

12.7.4.4 Vital Signs, Physical Findings and Other Observations Related to Safety

12.7.4.4.1 Prior and Concomitant Medication

Prior medication is defined as any medication that started and ended prior to Screening. Concomitant medication is defined as any medication starting on or after Screening. Medications that started prior to Screening and are ongoing at Screening are considered as concomitant.

All medications will be listed (Listing 15.3.6.4) by product using medicinal product and Anatomical Therapeutic and Chemical (ATC) codes (World Health Organization-Drug Dictionary Enhanced [WHO-DDE] Q1 2012 or later). A flag will be presented on the listing indicating whether the medication is prior or concomitant.

Prior and concomitant medications will be listed by randomization arm and will display original dates (no imputation). Concomitant medications will be summarized by randomization arm and separately by product use pattern category for the Safety population showing the number (%) of subjects who used the medication at least once by randomization arm and by ATC 1st and 2nd levels and medicinal product (Tables 15.2.6.24.*).

12.7.4.4.2 Physical Examination

Physical examination data recorded at V1, V3, V7, V10, and V16 will be listed by randomization arm (Listing 15.3.6.11). Subject's data with abnormal and abnormal clinically significant physical examination findings will be flagged. The number of subjects (%) with normal, abnormal and abnormal clinically significant results will be tabulated by body systems and randomization arm at Baseline, V7, V10 and V16 (Table 15.2.6.28). Summaries at V7, V10, and V16 will also be tabulated by product use pattern category (Table 15.2.6.29).

Body weight at V1, V3, and V5 to V16 Visits; and body height recorded at V1 will also be listed together with BMI (Listing 15.3.6.11). Listings will also include waist circumference assessments at V3, V7, V10, and V16.

Descriptive statistics of body weight, waist circumference, body height and BMI (BMI will also be categorized as shown in Section 4.1.3 "Categorical Variables") will be summarized at Baseline, V7, V10, V13 (weight only), and V16 by randomization arm (Table 15.2.6.30). Summaries for weight at V7, V10, V13, and V16 will also be tabulated by product use pattern category (Table 15.2.6.31).

12.7.4.4.3 Vital Signs

Systolic and diastolic blood pressure, pulse rate and respiratory rate measured during the study (V1, and V3 to V16) will be listed and summarized (Listing 15.3.6.9; Table 15.2.6.25). Vital signs assessments after Baseline will be listed and summarized together with change from Baseline by randomization arm.

12.7.4.4.4 Spirometry

Spirometry parameters assessed at V1, V3, V7, V10, and V16 (with and without salbutamol assessments at V1, V10, and V16) will be clinically interpreted (categories: normal, abnormal, abnormal clinically significant) and COPD staging will be evaluated [3].

The number and percentage of subjects with normal/abnormal/abnormal clinically significant results will be summarized together with changes in COPD categories by randomization arm (Listing 15.3.3.3; Table 15.2.6.27.1). Clinically significant shift tables by product use pattern category will be presented (Table 15.2.6.27.2).

12.7.4.4.5 Electrocardiogram

The ECG data will be obtained directly from the 12-lead ECG traces (*i.e.*, not centrally read). These data include the PR, QT, and QTcB intervals; QRS duration; heart rate; and normality evaluation (normal, abnormal, clinically relevant). In addition the QTcF value will be presented.

ECG data values and normality evaluations will be listed at V1, V10, and V16, together with changes from Baseline and shift in normality (Listing 15.3.6.10). ECG data from subjects which had significant clinical findings will be highlighted in listings. Clinical significant shift tables by product use pattern category will be presented (Table 15.2.6.26.2).

Descriptive statistics will be presented for ECG data at Baseline, V10, and V16 by randomization arm (Table 15.2.6.26.1). ECG data will be summarized together with changes from Baseline, and the number and percentage of subjects with normal/abnormal/abnormal clinically significant results.

12.7.4.4.6 Assessment of Cough

Cough symptoms will be assessed by means of the cough questionnaire, as described in Section 7.2.8 "Cough Questionnaire".

Descriptive analysis for change from Baseline will be presented for the VAS score and evaluating the level of cough bother at V7 and V10. For the calculation of the change from Baseline, missing VAS values will be imputed with zero when no cough was reported.

The responses to the individual items, including the VAS score, change from Baseline, 3 Likert scales measuring the intensity and frequency of cough, and the amount of sputum production will be listed and summarized using the number of subjects who reported to experience cough as denominator (Listing 15.3.3.5; Table 15.2.6.32). In addition, the maximum intensity, maximum frequency of cough, and maximum sputum production across all post-baseline visits will be summarized.

The number and percentage of subjects reporting a cough will be summarized by product use pattern categories by visit and overall across all post-baseline visits.

The answers to the open question related to any other important observation will be listed.

All descriptive summaries will be performed on the Safety population.

13 ANALYSIS AND REPORTING

13.1 Interim Analysis and Data Monitoring

No interim analysis is planned for the data collected during this study. Interim datasets and TFLs will be produced pre-database lock for programming QC purposes. In particular, TFLs will only be derived using pre-database lock data for disposition, baseline characteristics, and product use data; all other TFLs will be produced using dummy data generated for QC purposes. The data collected in the ZRHR-ERS-09-US study (the original study), however will be analyzed as planned and detailed in ZRHR-ERS-09-US study SAP v3.0 dated 07 March 2017.

A Clinical Research Associate ("Monitor") from [redacted] will be responsible for the monitoring of the study. Monitoring will be performed according to [redacted]'s standard operating procedures (SOPs) and as per the agreed monitoring plan with PMI.

The PI, or a designated member of the PI's staff, must be available during the monitoring visit to review the data and resolve any queries, and to allow direct access to the subject's records for source data verification.

All changes to the source data will have to be approved by the PI.

Upon finalization of this document, it will be specified which TFLs will be produced for Safety reporting, Topline results, Final Analysis, and Clinicaltrials.gov reporting purposes.

13.2 Safety Reporting

Statistical summaries required for safety reporting will be made available to PMI medical safety officer following database lock. These summaries are flagged in the TFLs list reported in Appendix 16.4 "Tables, Listings, and Figures".

13.3 Topline Results

Topline results, composed of key statistics and study results listings, will be made available to PMI management following database lock and prior to completion of the complete set of TFLs. The topline TFLs are listed in Appendix 16.4 "Tables, Listings, and Figures".

13.4 Final Analyses

Final analyses for this study will be performed only after database lock. A pre-analysis data review meeting will be held prior to database lock and completion of the final analyses. In addition, no database may be locked, randomization code unblinded, or analyses completed until the final version of this SAP has been approved.

Any post-hoc, additional exploratory analyses completed to support planned study analyses, which were not identified in this SAP, will be documented and reported as applicable. Any results from these unplanned analyses will also be clearly identified in the text of the CSR.

The list of all tables, figures and listings to be presented are included in the relevant sections of the SAP and reported in Appendix 16.4 "Tables, Listings, and Figures".

13.5 ClinicalTrials.gov Reporting

Statistical summaries which will be evaluated for publishing on the Clinicaltrials.gov website (NCT02649556) are listed in Appendix 16.4. "Tables, Listings, and Figures".

14 DATA PRESENTATION

Data presentation details are provided in a separate TFL mock shell document, based on the PMI style guide provided by PMI.

15 REFERENCES

1. ICH E9. Statistical principles for clinical trials. 1998; Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf (Accessed on 17 July 2014).

2. ICH E3. Structure and content of clinical study reports. 1995; Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf (Accessed on 17 July 2014).

3. Global initiative for Chronic Obstructive Lung Disease: GOLD. Pocket guide to COPD diagnosis, management, and prevention. 2013.

4. European Medicines Agency (EMA). Guideline on the development of medicinal products for the treatment of smoking (CHMP/EWP/369963/05). 2008.

5. Chemical Information Specialized Information Services. Nicotine [CAS No. 54-11-5]. Availble from: http://chemsisnlmnihgov/chemidplus/rn/54-11-5 (Accessed on 3 April 2014).

6. Chemical Information Specialized Information Services. Nicotine N-glucuronide [CAS No. 152306-59-7]. Available from: http://chemsisnlmnihgov/chemidplus/rn/152306-59-7 (Accessed on 3 April 2014).

7. Chemical Information Specialized Information Services. Cotinine [CAS No. 486-56-6]. Available from: http://chemsisnlmnihgov/chemidplus/rn/486-56-6 (Accessed on 3 April 2014).

8. Chemical Information Specialized Information Services. Cotinine N-glucuronide [CAS No. 139427-57-9]. Available from: http://chemsisnlmnihgov/chemidplus/rn/139427-57-9 (Accessed on 4 April 2014).

9. Chemical Information Specialized Information Services. Hydroxycotinine [CAS No. 34834-67-8]. Available from: http://chemsisnlmnihgov/chemidplus/rn/34834-67-8 (Accessed on 4 April 2014).

10. Chemical Information Specialized Information Services. Trans-3'-Hydroxycotinine-O-glucuronide [CAS No. 132929-88-5]. Availble from http://chemsisnlmnihgov/chemidplus/rn/132929-88-5 (Accessed on 4 April 2014).

11. Jacob P, 3rd, Yu L, Duan M, Ramos L, Yturralde O, Benowitz NL. Determination of the nicotine metabolites cotinine and trans-3'-hydroxycotinine in biologic fluids of smokers and non-smokers using liquid chromatography-tandem mass spectrometry: biomarkers for tobacco smoke exposure and for phenotyping cytochrome P450 2A6 activity. J Chromatogr B Analyt Technol Biomed Life Sci. 2011;879(3-4):267-76.

12. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict. 1991;86(9):1119-27.

13. Fagerstrom K, Russ C, Yu CR, Yunis C, Foulds J. The Fagerstrom Test for Nicotine Dependence as a predictor of smoking abstinence: a pooled analysis of varenicline clinical trial data. Nicotine Tob Res. 2012;14(12):1467-73.

14. Cappelleri JC, Bushmakin AG, Baker CL, Merikle E, Olufade AO, Gilbert DG. Confirmatory factor analyses and reliability of the modified cigarette evaluation questionnaire. Addict Behav. 2007;32:912–23.

15. DiClemente CC, Prochaska JO, Fairhurst S, Velicer WF, Rossi JS, Velasquez M. The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. J Consult Clin Psychol. 1991;56:295-304.

16. Velicer WF, Fava JL, Prochaska JO, Abrams DB, Emmons KM, Pierce J. Distribution of smokers by stage in three representative samples. Prev Med. 1995;24:401-11.

17. King BA, Hyland AJ, Borland R, McNeill A, Cummings KM. Socioeconomic variation in the prevalence, introduction, retention, and removal of smoke-free policies among smokers: findings from the International Tobacco Control (ITC) Four Country Survey. 2011.

18. Hailperin R. Best possible inequalities for the probability of a logical function of events. Am Math Mon. 1965;72(4):343-59.

19. Rüger B. Das maximale Signifikanzniveau des Tests: "Lehne H_0 ab, wenn k unter n gegebenen Tests zur Ablehnung führen". Metrika. 1978;25:171-8.

20. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. Stat Med. 2012;31(4):328-40.

21. Hu G, Cassano PA. Antioxidant nutrients and pulmonary function: the Third National Health and Nutrition Examination Survey (NHANES III). Am J Epidemiol. 2000;151(10):975-81.

22. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med. 1999;159(1):179-87.

23. Follmann DA. On the effect of treatment among would-be treatment compliers: an analysis of the multiple risk factor intervention trial. J Am Stat Assoc. 2000;95(452):1101-9.

24. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med. 2000;19(8):1059-79.

25. FDA (Food and Drug Administration). Implementation Specifications, reporting medical device problems (C54451/Medical_Device_Problem_Codes_FDA_CDRH). 2012; Available from: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/ReportingAdverseEvents/EventProblemCodes/ucm134761.htm (Accessed on 01 December 2014).

26. Greenland S, Lanes S, Jara M. Estimating effects from randomized trials with discontinuations: the need for intent-to-treat design and G-estimation. Clin Trials. 2008;5(1):5-13.

16 APPENDICES

16.1 Study Assessments

Refer to the ZRHR-ERS-09-US SAP for study assessment tables showing data collected in the original study.

Table A1 Study Assessments (separate table [Table A2] shown for 24 hour urine collections)

Visits Assessments	ZRHR-ERS-09-US	ZRHR-ERS-09-EXT-US Extended Exposure Period								Safety Follow-Up [g]
	V3 (Baseline)	V10		V11	V12	V13	V14	V15	V16	28 days
		ERS	EXT							
Study week (W)	-	W26		W30	W35	W39	W43	W48	W52	
Informed consent and additional ICFs for biobanking for ZRHR-ERS-09-EXT-US			●							
Inclusion / exclusion check for ZRHR-ERS-09-US	● V1+V3									
Inclusion / exclusion check for ZRHR-ERS-09-EXT-US			●							
Readiness to comply with study procedures and to use THS 2.2 for 26 additional weeks			●							
Enrollment of ZRHR-ERS- 09-US	●									
Enrollment of ZRHR-ERS-09-EXT-US			●							
Information on the risk of smoking, advice on smoking cessation (SC), and debriefing on THS 2.2	●	●		●	●	●	●	●	●	
Concomitant medication	●	●	●	●	●	●	●	●	●	
U: Pregnancy test (females)	●	●		●	●	●	●	●	●	
B/U: Clinical chemistry, hematology, urine analysis [a]	●	●							●	

Visits Assessments	ZRHR-ERS-09-US	ZRHR-ERS-09-EXT-US Extended Exposure Period								Safety Follow-Up [g]
	V3 (Baseline)	V10		V11	V12	V13	V14	V15	V16	28 days
ECG		●							●	
Vital signs [b]	●	●		●	●	●	●	●	●	
Waist circumference	●	●							●	
Weight and body mass index (BMI) [c]	●	●				●			●	
Physical examination	●	●							●	
Dispensing of THS 2.2	●		●	●	●	●	●	●		
CO breath test [d]	●	●							●	
U: BoExp in urine (Table A2)	●	●							●	
U: CVD clinical risk endpoints in urine (Table A2)	●	●							●	
B: CVD clinical risk endpoints (Table A3)	●	●							●	
B: BoExp in blood (Table A3)	●	●							●	
Pre-bronchodilator spirometry testing [e]		●							●	
Post-bronchodilator spirometry testing [e]	●	●							●	
Lung volume [e]	●	●							●	
Cough questionnaire	●	●							●	
MCEQ questionnaire	●	●							●	

Visits Assessments	ZRHR-ERS-09-US			ZRHR-ERS-09-EXT-US Extended Exposure Period						Safety Follow-Up [g]
	V3 (Baseline)	V10		V11	V12	V13	V14	V15	V16	28 days
Product use diary (daily) [f]	●		●	●	●	●	●	●	●	
FTND		●							●	
Intent to use THS 2.2 questionnaire (ITUQ) [h]		●							●	
AE/SAE recording [i]	●	●	●	●	●	●	●	●	●	●
U: Biobanking (Table A2) [j]	●	●							●	
B: Biobanking for BoExp and clinical risk endpoints (Table A3) [j]	●	●							●	
B: Biobanking for transcriptomics and lipidomics (Table A3) [j]	●	●							●	

Abbreviations: AE = Adverse event; B: Blood sample required; BMI = Body mass index; BoExp = Biomarkers of exposure; CC = Conventional cigarette(s); CO = Carbon monoxide; COHb = Carboxyhemoglobin; CVD = Cardiovascular disease; ECG = Electrocardiogram; FTND = Fagerström test for nicotine dependence; MCEQ = Modified cigarette evaluation questionnaire; SAE = Serious adverse event; THS = Tobacco Heating System; U = Urine sample required; VAS = Visual analog scale

[a] Safety laboratory parameters will be evaluated in at least 10 hours of fasting conditions.

Hematology: hematocrit, hemoglobin, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular volume, platelet count, red blood cell count, white blood cell count, and differential white blood cell count. Platelet count, white blood cell count from hematology will be evaluated as clinical risk endpoints. **Urine analysis**: pH, bilirubin, glucose, nitrite, red blood cell traces, protein, specify gravity. **Clinical chemistry**: albumin, total protein, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, blood urea nitrogen, creatinine, gamma-glutamyl transferase, glucose, lactate dehydrogenase, potassium, sodium, total bilirubin, direct bilirubin, total cholesterol, and triglycerides. White blood cell count (WBC) will also be evaluated as part of the "smokers' health profile".

[b] Systolic and diastolic blood pressure, pulse rate, and respiratory rate. Vital signs will be assessed after at least 5 minutes in supine position.

[c] Height will be transferred from the original

study.

[d] CO breath test: the test will be conducted once and should be done in conjunction with COHb blood where applicable.

[e] Lung function assessments must be performed in the following sequence:

- Pre-bronchodilator spirometry testing.
- Pre-bronchodilator lung volume using helium dilution (only at V10 and V16).
- Post-bronchodilator spirometry testing.

[f] Use of any tobacco/nicotine containing products will be captured in the ediary.

[g] Early termination procedures will be conducted in subjects who terminate from the study earlier.

[h] At V16, only subjects in the THS 2.2 arm will answer to the ITUQ.

[i] During the safety follow-up period, spontaneous reporting of new AEs/SAEs by the subject will be done.

[j] Once at every visit. Samples will only be taken if additional ICF for sample biobanking is signed by the subject.

All assessments done at V10 in the context of the original study do not need to be repeated for the extension study.

Table A2 Schedule for 24-hour Urine Collection Assessments

	ZRHR-ERS-09-US		ZRHR-ERS-09-EXT-US
	V3 (Baseline)	V10	V16
	ERS	EXT	
Exposure			
BoExp in urine [a]	•	•	•
CVD Clinical Risk Endpoints			
11-DTXB_2, 8-epi-PG$F_{2\alpha}$, albumin	•	•	•
Others			
Creatinine	•	•	•
Biobanking for BoExp and clinical risk endpoints [b]	•	•	•

Abbreviations: BoExp = Biomarker(s) of exposure; CVD = Cardiovascular disease; 8-epi-PG$F_{2\alpha}$ = 8-epi-prostaglandine $F_{2\alpha}$; 11-DTXB_2 = 11- dehydrothromboxane B_2.

[a] 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (total NNAL), total N-nitrosonornicotine (total NNN), nicotine equivalents (Neq).

[b] Samples will only be taken if additional ICF for biobanking is signed by the subject

All assessments done at V10 in the context of the original study do not need to be repeated for the extension study.

Table A3 Schedule for Blood/Serum/Plasma Assessments (except Hematology and Clinical Chemistry)

	ZRHR-ERS-09-US		ZRHR-ERS-09-EXT-US
	V3 (Baseline)	V10	V16
	ERS	EXT	
Blood			
CVD clinical risk endpoints (COHb and HbA1c)[a]	•	•	•
Biobanking [b]	•	•	•
Plasma			
CVD clinical risk endpoints (fibrinogen and homocysteine)	•	•	•
Biobanking [b]	•	•	•
BoExp: nicotine and cotinine	•	•	•
Serum			
CVD clinical risk endpoints [c]	•	•	•
Biobanking [b]	•	•	•

Abbreviations: BoExp = Biomarker(s) of exposure; CVD = Cardiovascular disease; COHb = Carboxyhemoglobin; HbA1c = Hemoglobin glycosylated A1c

[a] WBC and platelet count will be taken from the laboratory safety parameters.

[b] Samples must only be taken if additional ICF for biobanking is signed by the subject.

[c] sICAM-1, MPO, LDL-C, HDL-C, hs-CRP, Apo A1 and Apo B.

All assessments done at V10 in the context of the original study do not need to be repeated for the extension study.

16.2 Statistical Considerations for the Definition of Analysis Populations

In pivotal parallel group pharmaceutical therapy superiority studies the general approach is to analyze the intention-to-treat (ITT) population. In such studies the product efficacy and the product compliance are generally highly correlated as the product effects may be detectable by the subject (*e.g.*, relief from symptoms). The drivers for product use (ultimately product compliance) for tobacco products are not the same as in a pharmaceutical setting. Smokers decide to smoke, despite the known risks and adverse health effects, based on social, cultural, personal, and psychological factors. In addition, classical ITT analyses would compare subjects based on their assigned (intended) exposure, in order to preserve randomization, without regard to the subsequent compliance. As a result, estimates of exposure effects from these comparisons would be confounded with compliance, and can be seriously underestimated when noncompliance is high [26]. Therefore the ITT approach may not be optimal in randomized studies evaluating modified risk tobacco products, as these products are targeting risk reduction which may not be the main factor influencing the choice of tobacco products used by the subjects and the level of noncompliance to the product (*e.g.*, Dual-use) may not be negligible.

The aim of this study is to demonstrate the real-world effect that could be achieved by the exposure to THS 2.2, in terms of favorable changes in the 'smokers 'health profile' endpoints. In this perspective, the following considerations were taken into account in order to measure effect in this study:

- The primary analysis is designed to assess the effect in all subjects as per actual use of study products.

- The subjects are allowed to use the THS 2.2 product in a truly ad libitum setting (no product use restrictions are made in the study protocol).

To support the interpretation of the results, the analysis of the primary objective is conducted on the PP population and on the FAS-AR (see Section 12.7.1.3), and including a sensitivity analysis conducted in the context of the exploratory exposure-response model (see Section 12.7.3.2).

16.3 Data Handling and Derivation of Product Use Pattern Categories

Only available data collected since the first day after randomization (Day 2) up until V16 (Day V16 excluded) will be considered for product use categorization for product use pattern categorization.

It is expected that any duplicate product use record (i.e. more records on the same study day) is identified during data review and flagged for the exclusion from the analysis.

16.3.1 Preliminary Data Handling

The following data handling will be performed on the two 6-month analysis periods separately.

The following data of the product use diary will not be considered for product use categorization and will be set to missing:

- Records of days with reported CC use for subjects in the THS 2.2 arm who report CC use for less than 5 days in total.
- Records of days with reported THS use for subjects in the CC arm who report THS 2.2 use for less than 5 days in total
- Records of days with outliers in any of the product use, identified as those values which are above 60 units and also greater than median+3*IQRange, with IQRange being the inter-quartile range of the specific product use data distribution.

Missing dates for Visits 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 and 16 will be imputed for the calculation of product use exposure within periods by selecting the middle date between two non-missing visits. For example, if Visit 6 and Visit 7 are missing, then take the number of days between Visit 5 and Visit 8 and divide by 3 and round to the nearest day. Add this resulting number of days to Visit 5 to determine the Visit 6 imputed study day and then add this resulting number of days to Visit 6 to determine the Visit 7 imputed study day. For an early terminated subject, add 30 days after the last visit to determine each successive visit's study day.

Missing data will be interpolated using a moving average over a 15-day window (± 7 days) centered on the missing data. For example, someone missing THS data on Day 10 would have the average of non-missing values between Days 3 to 17 interpolated for Day 10.

16.3.2 Derivation of Actual Product Use Categories

Per-Day product use categories will be derived for each study day as indicated in Table 19, and the percent of days (over the whole analysis period) the subjects fall in each category will be computed.

Table 19: Per Day Actual Product Use Pattern Categories

Category Label	Definition
THS-use	nTHS + nCC ≥ 1 and nECig/(nTHS+nCC+nECig+nOtherTob) <50% and nOtherTob/(nTHS+nCC+nECig+nOtherTob) <50% and nTHS/(nTHS + nCC) ≥ 70%
Dual-use	nTHS + nCC ≥ 1 and nECig/(nTHS+nCC+nECig+nOtherTob) <50% and nOtherTob/(nTHS+nCC+nECig+nOtherTob) <50% and 1% ≤ nTHS/(nTHS + nCC) < 70%
CC-use	nTHS + nCC ≥ 1 and nECig/(nTHS+nCC+nECig+nOtherTob) <50% and nOtherTob/(nTHS+nCC+nECig+nOtherTob) <50% and nTHS/(nTHS + nCC) < 1%
ECig-use	1) nTHS + nCC ≥ 1 and nECig ≥ nOtherTob and [nECig/(nTHS+nCC+nECig+nOtherTob) ≥50% or nOtherTob/(nTHS+nCC+nECig+nOtherTob) ≥50%]; or 2) nTHS + nCC < 1 and nECig ≥ nOtherTob and (nEcig ≥ 1 or nOtherTob ≥ 1)
OtherTobacco-use	1) nTHS + nCC ≥ 1 and nECig < nOtherTob and [nECig/(nTHS+nCC+nECig+nOtherTob) ≥50% or nOtherTob/(nTHS+nCC+nECig+nOtherTob) ≥50%]; or 2) nTHS + nCC < 1 and nECig < nOtherTob and (nEcig ≥ 1 or nOtherTob ≥ 1)
Abstainer	nTHS + nCC < 1 and nECig < 1 and nOtherTob < 1
Missing	Missing product use data

Note: nTHS, nCC, nECig, nOtherTob indicates the quantity of daily product use reported respectively for THS 2.2, CC, E-cigarette, and other tobacco..

The product use categories for the 3-month analysis periods (]V4, V7],]V7, V10[, [V10, V13],]V13, V16[) are derived following the same rules defined for per-day categories in Table 19, but using the average product use over the 3-month period. In addition:

- In order to factor in the higher relevance of later product use in the period for their impact on the endpoints, a weighted average approach is adopted, using the day number as weighting factor. The day numbers will be relative to the analysis period (i.e. weights ranging between 1 and the last day number of the period, normalized to

the total sum of the weights).

- THS-use or CC-use categories will be assigned to the analysis period when consistently observed on a per-day basis over the same period. In particular the frequency of THS-use and of CC-use per-day categories is calculated using the same weighting scheme previously adopted to compute the average product use over the entire 3-month period. THS-use and CC-use product use categories would be turned into Dual-use category if the weighted percent of days in respectively THS-use and CC-use is less than 50% during the analysis period.
- Subjects in the THS-use category will be flagged as Exclusive-use if nTHS/(nTHS + nCC) ≥ 95%. Subjects in CC-use will be flagged as Exclusive-use if nTHS/(nTHS + nCC) ≤ 5%.
- Setting to Missing the product use categories if, after the interpolation, more than one third of the days (e.g. 30 days for the 3-month periods) of missing product use category during the analysis period. For the sensitivity analysis on missing product use data (see Section 12.7.1.3.4), the non-imputed product use category will be defined based on the available data (i.e. categories will be set to Missing only if no data is recorded for the analysis period).

The product use categories for the six-month]V4,V10[and [V10,V16[analysis periods are derived following the same approach used for the 3-month periods, with the following differences:

- A weighting scheme for the calculation of the average of the product use over the 6-month period is adopted so to reflect the increasing relevance of the product use over time for their association with the value of the endpoints at the end ot the period and especially during the last 3 months. To this aim, before being normalized the weights are set as follows:
    - Set to 1.0 for the]V4, V7] and [V10, V13] parts of each 6-month analysis period
    - Range between 1.0 up until the last study day relative to the]V7, V10[and]V13, V16[time window, for the last 3-months of each 6-month analysis period.
- For each 6-month period, subjects having product use categories different between the two 3-months period components will move to a category with label “Switcher” if:

- Category is THS-use in one 3-month sub-period, and different from Dual-use (either CC-use or Other-use) in the other sub-period
- Category is CC-use or ECig-use or OtherTobacco-use or Abstainer in only one of the 3-month sub-periods

- For analysis purposes, an over-arching category with label "Other-use" will be used to incorporate all product use categories different from THS-use, Dual-use, or CC-use. The detailed categories will be used only for the descriptive summaries of the product use over time.
- Setting to Missing the category if at least one of the categories derived for the two 3-month sub-periods is missing, apart from:
  - If]V7,V10[category is Missing when the subjects early terminate after V7 and category of]V4, V7] is not missing, the latter will be assigned to]V4, V10[.
  - If]V13, V16[category is Missing when the subjects early terminate after V10 and category of [V10, V13] is not missing, the latter will be assigned to [V10, V16[.

  For the sensitivity analysis on missing product use data (see Section 12.7.1.3.4), the category will be set to Missing only if the categories of both the two 3-month sub-periods are set to missing, otherwise the category will be set to the non-missing category of the two 3-month sub-periods.
- The product use categories for the 12-month randomized period are derived combining the product use categories derived for the two 6-month periods as reported in Table 20.

Table 20: Overall Product Use Pattern Categories

1st/2nd 6-month Product use	THS-use	CC-use	Dual-use	ECig-use	OtherTobacco-use	Abstainer	Switcher
THS-use	THS-use	Switcher	Dual-use	Switcher	Switcher	Switcher	Switcher
CC-use	Switcher	CC-use	Switcher	Switcher	Switcher	Switcher	Switcher
Dual-use	Dual-use	Switcher	Dual-use	Switcher	Switcher	Switcher	Switcher
ECig-use	Switcher	Switcher	Switcher	ECig-use	Switcher	Switcher	Switcher
OtherTobacco-use	Switcher	Switcher	Switcher	Switcher	OtherTobacco-use	Switcher	Switcher

Abstainer	Switcher	Switcher	Switcher	Switcher	Switcher	Abstainer	Switcher
Switcher	Switcher	Switcher	Switcher	Switcher	Switcher	Switcher	Switcher

Row and columns indicates the product use categories during the first and second six months analysis periods

- For subjects who do not enroll into the extension study, or for subjects who early terminate after V10 and category is missing for the period [V10, V16[, the product use category from]V4,V10[will be assigned for the product use category of the 12-month randomized period.
- For analysis purposes, an over-arching category with label “Other-use” will be used to incorporate all product use categories different from THS-use, Dual-use, or CC-use. The detailed categories will be used only for the descriptive summaries of the product use over time.
- Setting to Missing the category if at least one of the categories derived for the two 6-month analysis periods is missing, apart from the subjects who early terminate after V10 and category of first 6-month period is not missing.
  For the sensitivity analysis on missing product use data (see Section 12.7.1.3.4), the category will be set to Missing only if the categories of both the two 6-month periods are set to missing, otherwise the category will be set to the non-missing category of the two 6-month periods.

16.4 Tables, Listings, and Figures

The table below reports all the TFLs and flags those required for topline review, safety reporting, and Clinical Trial.gov reporting.

Tables

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.1.1	Summary of Subject Disposition – Screened Subjects	Y		Y
Table 15.2.1.2	Summary of Reasons for Discontinuations – Enrolled Subjects			
Table 15.2.1.3.1	Summary of Protocol Deviations – Randomized Population			
Table 15.2.1.3.2	Analysis Sets and Reasons for Exclusions from Analyses – Screened Subjects			
Table 15.2.1.4.1	Summary of Demographics and Other Baseline Characteristics by 12-Month Product Use Categories – FAS-EX	Y		Y
Table 15.2.1.4.1.1	Summary of Demographics and Other Baseline Characteristics by 12-Month Product Use Categories and by Sex – FAS-EX			
Table 15.2.1.4.1.2	Summary of Demographics and Other Baseline Characteristics by 6-Month Product Use Categories – FAS-EX			Y

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.1.4.1.3	Summary of Demographics and Other Baseline Characteristics by 6-Month Product Use Categories – FAS-EX Subjects who Enrolled into Extension Study			
Table 15.2.1.4.1.4	Summary of Demographics and Other Baseline Characteristics by 12-Month Product Use Categories – FAS-EX Subjects who Enrolled into Extension Study			
Table 15.2.1.4.2.1	Summary of Demographics and Other Baseline Characteristics – Safety Population			
Table 15.2.1.4.2.2	Summary of Demographics and Other Baseline Characteristics – PP Population			
Table 15.2.1.4.2.3	Summary of Demographics and Other Baseline Characteristics – FAS-AR			
Table 15.2.1.4.3.1	Summary of Baseline Covariates Comparability for Smokers' Health Profile Analysis by 6-Month Product Use Categories – FAS-EX	Y		
Table 15.2.1.4.3.2	Summary of Baseline Covariates Comparability for Smokers' Health Profile Analysis by 12-Month Product Use Categories – FAS-EX			
Table 15.2.1.4.4	Summary of Baseline Covariates Comparability by Enrollment to			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
	Extension Study – FAS-EX			
Table 15.2.1.5	Summary of Medical History – Safety Population			
Table 15.2.1.6	Summary of Concomitant Diseases – Safety Population			
Table 15.2.2.1.1	Summary of Average Daily Product Use in Randomized Exposure Period by 12-Month Product Use Categories – FAS-EX			
Table 15.2.2.1.1.1	Summary of Average Daily Product Use in Randomized Exposure Period by 12-Month Other-use Detailed Categories – FAS-EX			
Table 15.2.2.1.1.2	Summary of Average Daily Product Use in Randomized Exposure Period by 12-Month Product Use Categories and by Sex - FAS-EX			
Table 15.2.2.1.1.3	Summary of Average Daily Product Use in Randomized Exposure Period by 12-Month Product Use Categories and by Consumption Level - FAS-EX			
Table 15.2.2.1.1.4	Summary of Average Daily Product Use in Randomized Exposure Period by 6-Month Product Use Categories – FAS-EX			
Table 15.2.2.1.2	Summary of Average Daily Product Use in Randomized Exposure Period – PP Population			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.2.1.3	Summary of Average Daily Product Use in Randomized Exposure Period by Product Use – Safety Population			
Table 15.2.2.1.3.1	Summary of Average Daily Product Use by Arm – Safety Population			
Table 15.2.3.1	Covariates included in Analysis of Smokers' Health Profile	Y		
Table 15.2.3.1.1	Analysis of Smokers' Health Profile – FAS-EX	Y		Y
Table 15.2.3.1.2	Missing Product Use Sensitivity Analysis of Smokers' Health Profile – FAS-EX			
Table 15.2.3.1.3	Sensitivity Analysis of Smokers' Health Profile for Subjects Enrolled into Extension Study – FAS-EX			
Table 15.2.3.1.4	Multiple Imputation Sensitivity Analysis of Smokers' Health Profile – FAS-EX			
Table 15.2.3.1.5	Sensitivity Analysis of Smokers' Health Profile Derived Endpoints Including LLOQ Values in Denominator – FAS-EX			
Table 15.2.3.2.1	Supportive Analysis of Smokers' Health Profile – FAS-EX	Y		

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.3.2.2	Supportive Analysis of Smokers' Health Profile – PP Population			
Table 15.2.3.2.3	Supportive Analysis of Smokers' Health Profile – FAS-AR			
Table 15.2.3.3.1	Secondary Analysis of Biomarkers of Exposure – FAS-EX	Y		
Table 15.2.3.3.1.1	Sensitivity Analysis of Biomarkers of Exposure Including LLOQ Values in Denominator – FAS-EX			
Table 15.2.3.3.2	Secondary Analysis of Biomarkers of Exposure – PP Population			
Table 15.2.3.3.3	Secondary Analysis of Biomarkers of Exposure – FAS-AR			
Table 15.2.4.1.1	Descriptive Statistics of Smokers' Health Profile - FAS-EX			
Table 15.2.4.1.1.1	Descriptive Statistics of Smokers' Health Profile by Sex- FAS-EX			
Table 15.2.4.1.1.2	Descriptive Statistics of Smokers' Health Profile by Consumption Level - FAS-EX			
Table 15.2.4.1.2	Descriptive Statistics of Smokers' Health Profile – PP Population			
Table 15.2.4.1.3	Descriptive Statistics of Smokers' Health Profile – FAS-AR			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.4.2.1	Descriptive Statistics of Biomarkers of Exposure - FAS-EX			
Table 15.2.4.2.2	Descriptive Statistics of Biomarkers of Exposure – PP Population			
Table 15.2.4.2.3	Descriptive Statistics of Biomarkers of Exposure - FAS-AR			
Table 15.2.4.3.1	Descriptive Statistics of Neq, Plasma Nicotine and Cotinine Concentrations – FAS-EX			
Table 15.2.4.3.2	Descriptive Statistics of Neq, Plasma Nicotine and Cotinine Concentrations – PP Population			
Table 15.2.4.3.3	Descriptive Statistics of Neq, Plasma Nicotine and Cotinine Concentrations – FAS-AR			
Table 15.2.4.4.1	Analysis of Neq, Plasma Nicotine and Cotinine Concentrations – FAS-EX			
Table 15.2.4.4.1.1	Sensitivity Analysis of Neq Including LLOQ Values in Denominator – FAS-EX			
Table 15.2.4.4.2	Analysis of Neq, Plasma Nicotine and Cotinine Concentrations – PP Population			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.4.4.3	Analysis of Neq, Plasma Nicotine and Cotinine Concentrations – FAS-AR			
Table 15.2.4.5.1	Descriptive Statistics of Additional Lung Function Markers - FAS-EX			
Table 15.2.4.5.2	Descriptive Statistics of Additional Lung Function Markers – PP Population			
Table 15.2.4.5.3	Descriptive Statistics of Additional Lung Function Markers – FAS-AR			
Table 15.2.4.6.1	Analysis of Additional Lung Function Markers – FAS-EX	Y		
Table 15.2.4.6.2	Analysis of Additional Lung Function Markers – PP Population			
Table 15.2.4.6.3	Analysis of Additional Lung Function Markers – FAS-AR			
Table 15.2.4.7.1	Descriptive Statistics of Additional Cardiovascular Disease Markers by Product Use Pattern - FAS-EX			
Table 15.2.4.7.2	Descriptive Statistics of Additional Cardiovascular Disease Markers – PP Population			
Table 15.2.4.7.3	Descriptive Statistics of Additional Cardiovascular Disease Markers			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
	– FAS-AR			
Table 15.2.4.8.1	Analysis of Additional Cardiovascular Disease Markers – FAS-EX	Y		
Table 15.2.4.8.1.1	Sensitivity Analysis of Additional Cardiovascular Disease Markers Including LLOQ Values in Denominator – FAS-EX			
Table 15.2.4.8.2	Analysis of Additional Cardiovascular Disease Markers – PP Population			
Table 15.2.4.8.3	Analysis of Additional Cardiovascular Disease Markers – FAS-AR			
Table 15.2.4.9.1	Summary of Cough Assessment Over Study – FAS-EX			
Table 15.2.4.9.2	Summary of Cough Assessment Over Study – PP Population			
Table 15.2.4.10	Analysis of Proportion of Subjects Reporting Need to Cough Over Study – FAS-EX			
Table 15.2.4.11.1	Descriptive Statistics of MCEQ Subscales – FAS-EX			
Table 15.2.4.11.2	Descriptive Statistics of MCEQ Subscales – PP Population			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.4.12.1	Analysis of MCEQ Subscales – FAS-EX			
Table 15.2.4.12.2	Analysis of MCEQ Subscales – PP Population			
Table 15.2.4.13.1	Descriptive Statistics of FTND Results – FAS-EX			
Table 15.2.4.13.2	Descriptive Statistics of FTND Results – PP Population			
Table 15.2.4.14.1	Descriptive Statistics of Intent to Use THS 2.2 Questionnaire – FAS-EX			
Table 15.2.4.14.2	Descriptive Statistics of Intention to Use THS 2.2 Questionnaire – PP Population			
Table 15.2.4.15	Descriptive Statistics of Distribution of Dual Use over Randomized Exposure Period – FAS-EX			
Table 15.2.4.16	Analysis of Final Propensity Model – FAS-AR			
Table 15.2.4.17	Analysis of Exposure-Response Relationship of Dual Use – FAS-AR			
Table 15.2.5.1	Summary of Product Use Pattern – FAS-EX	Y		

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.5.1.1	Summary of Product Use Pattern by Sex - FAS-EX			
Table 15.2.5.1.2	Summary of Product Use Pattern by Consumption Level - FAS-EX			
Table 15.2.5.2	Summary of Product Use Pattern - Safety Population			
Table 15.2.5.3	Summary of Product Use Pattern - FAS-AR	Y		
Table 15.2.5.4	Summary of Detailed Product Use Pattern - FAS-EX			
Table 15.2.6.1	Summary of Adverse Events – Safety Population		Y	
Table 15.2.6.2	Summary of Adverse Events by Product Use Pattern Category – Safety Population	Y	Y	
Table 15.2.6.3	Summary of Adverse Events by System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.3.1	Summary of Monthly Adverse Events by System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.4	Summary of Adverse Events by Product Use Pattern Category, System Organ Class and Preferred Term – Safety Population	Y	Y	

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.6.4.1	Summary of Frequent Adverse Events (Incidence >5%) by System Organ Class and Preferred Term and Product Use Category– Safety Population		Y	Y
Table 15.2.6.5	Summary of Adverse Events by System Organ Class, Preferred Term and Relationship to Study Product Exposure and Expectedness – Safety Population		Y	
Table 15.2.6.6	Summary of Adverse Events by Product Use Pattern Category, System Organ Class, Preferred Term and Relationship to Study Product Exposure and Expectedness – Safety Population	Y	Y	
Table 15.2.6.7	Summary of Adverse Events Leading to Study Product Discontinuation, Interruption, or Reduction by System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.8	Summary of Adverse Events Leading to Study Product Discontinuation, Interruption, or Reduction by Product Use Pattern Category, System Organ Class, and Preferred Term – Safety Population		Y	
Table 15.2.6.9	Summary of Adverse Events Related to Study Procedure by System Organ Class and Preferred Term – Safety Population		Y	

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.6.10	Summary of Adverse Events Related to Study Procedure by Product Use Pattern Category, System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.11	Summary of Adverse Events by System Organ Class, Preferred Term and Severity – Safety Population		Y	
Table 15.2.6.12	Summary of Adverse Events by Product Use Pattern Category, System Organ Class, Preferred Term and Severity – Safety Population		Y	
Table 15.2.6.13	Summary of Serious Adverse Events by System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.14	Summary of Serious Adverse Events by Product Use Pattern Category, System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.15	Summary of Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term – Safety Population		Y	
Table 15.2.6.16	Summary of Adverse Events Leading to Study Discontinuation by Product Use Pattern Category, System Organ Class and Preferred Term – Safety Population		Y	

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.6.17	Summary of THS 2.2 Device Events – Safety Population		Y	
Table 15.2.6.18	Summary of Clinical Chemistry Parameters – Safety Population		Y	
Table 15.2.6.19	Summary of Clinical Chemistry Parameters by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.20	Summary of Hematology Parameters – Safety Population		Y	
Table 15.2.6.21	Summary of Hematology Parameters by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.22	Summary of Urinalysis Parameters – Safety Population		Y	
Table 15.2.6.23	Summary of Urinalysis Parameters by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.24.1	Summary of Concomitant Medication by Anatomical Therapeutic Classes (ATC) 1 and 2 – Safety Population		Y	
Table 15.2.6.24.2	Summary of Concomitant Medication by Anatomical Therapeutic Classes (ATC) 1 and 2 and Product Use Pattern Category – Safety Population		Y	

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.6.24.3	Summary of Concomitant Medication by Medicinal Product – Safety Population		Y	
Table 15.2.6.24.4	Summary of Concomitant Medication by Medicinal Product and Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.25	Summary of Supine Vital Signs – Safety Population		Y	
Table 15.2.6.26.1	Summary of ECG Results – Safety Population		Y	
Table 15.2.6.26.2	Shift from Baseline of ECG Results by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.27.1	Summary of Spirometry Results – Safety Population		Y	
Table 15.2.6.27.2	Shift from Baseline of Spirometry by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.28	Summary of Physical Examination of Body Systems – Safety Population		Y	
Table 15.2.6.29	Summary of Physical Examination of Body Systems by Product Use Pattern Category – Safety Population		Y	

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Table 15.2.6.30	Summary of Weight, Waist Circumference and BMI Results – Safety Population		Y	
Table 15.2.6.31	Summary of Weight Results by Product Use Pattern Category – Safety Population		Y	
Table 15.2.6.32	Summary of Cough Assessment Over Study – Safety Population		Y	
Figure 15.1.1.1	Forest Plot of Supportive Analysis of Smokers' Health Profile – FAS-EX			
Figure 15.1.1.2	Forest Plot of Supportive Analysis of Smokers' Health Profile – PP Population.			
Figure 15.1.1.3	Forest Plot of Supportive Analysis of Smokers' Health Profile – FAS-AR			
Figure 15.1.2.1	Histogram of Predicted Propensity Scores of Exposure-Response Relationship – FAS-AR			
Figure 15.1.2.2	Forest Plot of Statistical Analysis of Exposure-Response Relationship – FAS-AR			
Listing 15.3.1.1	Listing of Subject Disposition			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Listing 15.3.1.1.1	Listing of Visits			
Listing 15.3.1.2	Listing of Unmet Inclusion / Exclusion Criteria of Extension Study			
Listing 15.3.1.3	Listing of Questions on Smoking History/Habits			
Listing 15.3.1.4	Listing of Prochaska “Stage of Change” Questionnaire Results			
Listing 15.3.1.5	Listing of Product Preference			
Listing 15.3.1.6	Listing of Urine Pregnancy Results			
Listing 15.3.1.7	Listing of Demographics			
Listing 15.3.1.8	Listing of Assignment to Analysis Sets			
Listing 15.3.1.9	Listing of Medical History and Concomitant Diseases			
Listing 15.3.1.10	Listing of Protocol Deviations			
Listing 15.3.1.11	Listing of Socio-Economic Questionaire Results			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Listing 15.3.1.12	Listing of Lifestyle Assessment Results			
Listing 15.3.1.13	Listing of Attempt to Quit Smoking			
Listing 15.3.2.1	Listing of Product Usage			
Listing 15.3.2.1.1	Listing of Product Use Categories			
Listing 15.3.2.1.2	Listing of Derived Product Use Summaries			
Listing 15.3.2.2	Listing of THS 2.2 Device Details			
Listing 15.3.2.3	Listing of THS 2.2 Product Accountability			
Listing 15.3.2.4	Listing of Subjects and Observations Excluded from Analysis Populations			
Listing 15.3.3.1	Listing of Urinary Biomarkers			
Listing 15.3.3.2	Listing of Blood, Plasma, Serum, and Exhaled Air Biomarkers			
Listing 15.3.3.3	Listing of Full Lung Function Data and Changes from Baseline			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Listing 15.3.3.3.1	Listing of Not Done Biomarker and Lung Function Parameters			
Listing 15.3.3.4	Listing of Plasma Nicotine and Cotinine Concentrations			
Listing 15.3.3.5	Listing of Cough Assessment Results			
Listing 15.3.4.1	Listing of MCEQ Questionnaire Results			
Listing 15.3.5.1	Listing of FTND Results			
Listing 15.3.5.2	Listing of Intent to Use THS 2.2 Questionnaire Results			
Listing 15.3.6.1	Listing of Adverse Events			
Listing 15.3.6.2	Listing of Serious Adverse Events			
Listing 15.3.6.3	Listing of Adverse Events Leading to Study Discontinuation			
Listing 15.3.6.4	Listing of Prior and Concomitant Medication			
Listing 15.3.6.5	Listing of THS 2.2 Device Events and Malfunctions			

Table No.	Title	Topline (Y/N)	Safety (Y/N)	Clinical Trial.gov (Y/N)
Listing 15.3.6.6	Listing of Clinical Chemistry Data, Shift, Changes from Baseline and CTCAE grades			
Listing 15.3.6.7	Listing of Hematology Data, Shift, Changes from Baseline and CTCAE grades			
Listing 15.3.6.8	Listing of Urinalysis Data, Shift, Changes from Baseline and CTCAE grades			
Listing 15.3.6.9	Listing of Vital Signs Data and Changes from Baseline			
Listing 15.3.6.10	Listing of ECG Data and Changes from Baseline			
Listing 15.3.6.11	Listing of Physical Examination Findings, Shift and Changes from Baseline			