PMI RESEARCH & DEVELOPMENT

Statistical Analysis Plan

Study Title:	A randomized, controlled, open-label, 4-arm parallel group study to evaluate the effect of switching from cigarette smoking to the use of IQOS in healthy adult current smokers on exercise capacity and trainability
Study Number:	P1-EXC-01-EU
Product Name:	IQOS
Sponsor:	Philip Morris Products S.A. Quai Jeanrenaud 5 2000 Neuchâtel, Switzerland
Version:	Final 2.0
Date:	10 Mar 2020

STATISTICAL ANALYSIS PLAN APPROVAL SIGNATURES

When this page is signed 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] Senior Statistician [redacted] Philip Morris Products S.A.	[redacted] Senior Clinical Scientist [redacted] Philip Morris Products S.A.
[redacted] Senior Officer Medical Safety [redacted] Philip Morris Products S.A.	[redacted] Manager Biostatistics [redacted] Philip Morris Products S.A.

TABLE OF CONTENTS

1 Introduction....5
1.1 Revision History....5
2 Abbreviation of Terms....6
3 STUDY OBJECTIVES AND ENDPOINTS....8
3.1 Objectives and Endpoints....8
3.2 Exploratory Endpoints....11
3.3 Study Hypotheses and Evaluation Criteria....12
4 Investigational Plan....13
4.1 Study Design....13
4.2 Selection of Study Population....16
4.3 Product Allocation and Blinding....20
5 Derived and Computed Variables....21
5.1 Urinary Biomarkers....21
5.2 Blood composition as determined by CO re-breathing method....22
5.3 Questionnaires....23
5.4 Categorical Variables....27
5.5 Product Use....27
6 Sample Size Justification....29
7 Changes from the Protocol Specified Statistical Analyses....30
8 Analysis Sets....30
8.1 Full Analysis Set....30
8.2 Safety Set....30
8.3 As Exposed Exercise Compliant Set....31
8.4 As Exposed Biochemically and Exercise Compliant Set....32
8.5 Protocol Deviations....32
9 Planned Statistical Methods....33
9.1 General Considerations....33
9.2 Disposition of Subjects....37

9.3 Demographic and Other Baseline Characteristics....38
9.4 Measurement of Product Compliance....38
9.5 Extent of Exposure (Product Consumption)....38
9.6 Planned Statistical Analyses....39
10 Analyses and Reporting....56
10.1 Interim Analyses and Data Monitoring....56
10.2 Safety Reporting....56
10.3 Topline Results....56
10.4 Final Analyses....57
10.5 ClinicalTrials.Gov Reporting....57
11 DATA PRESENTATION....57
12 References....57
13 APPENDICES....58
13.1 Study Assessments....58
13.2 Tables, Figures & Listings....64

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes the methodology and considerations of the planned analyses and lists the Tables, Figures and Listings (TFLs) for this study. A detailed description of the TFLs will be provided in a separate TFL shell document. Any changes to the TFL shell numbering or to the title of a TFL will not require an amendment to the SAP.

This SAP will be finalized prior to the lock of the clinical database. 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"
- ICH E3 guideline entitled, "Guidance for Industry: Structure and Content of Clinical Study Reports"
- Electronic case report forms (eCRF) final version 4.0 (dated 28 Aug 2019)
- Tables, Listings, and Figures style guide (PMI-RRP-FOR-112420) (dated 26 Jun 2018)

1.1 Revision History

Table should only include approved version and needs to be completed starting with version 1, the initial version of the document.

Version	Date of Revision	Revision
1.0	06 Dec 2019	Initial Version
2.0	09 Mar 2020	• Updated wording in formula for calculation of Nicotine Equivalents in 5.1.1. • Added IPAQ scoring algorithm in 5.3.5 • The calculation of product use categories is simplified to only be done for Week 1 and from V3 – V43 (Overall) or visit before early termination where applicable. For subjects with insufficient product use data they will be classified under the "Other" group per section 5.5.1. • The definition of completion of a bike training is expanded to include cases where at least 20 minutes of training was done and also those which were done but without data due to technical issues. • Added 9.1.7 to describe how to handle data which is implausible.

Version	Date of Revision	Revision
		• Descriptive statistics for $VO2_{max}$, exercise capacity and other related endpoints will now include an acute and training baseline that is based on product use categories calculated at Week 1 and Overall, respectively, this is related to third bullet point above. • Clarified the analysis in 9.6.1.1.2 for exercise capacity to include or exclude subjects which did not have the correct work load performed. Added a sensitivity analysis to assess the impact when all subjects including those with the wrong workloads are analysed. • Removed some boxplots for VO_2max and exercise capacity based on subgroups. • Added demographics for AEBECS • Added tables for analysis sets and reason for exclusion from analysis sets and number of subjects in each analysis set. • Added tables for exercise capacity, biological health markers and biomarkers of exposure for AEBECS. • Removed some redundant listings and updated listing numbers. • Fixed typos throughout document.

2 ABBREVIATION OF TERMS

AE	Adverse event
AEBECS	As Exposed Biochemically and Exercise Compliant Set
AEECS	As Exposed Exercise Compliant Set
BMI	Body mass index
BoExp	Biomarker of exposure
Bpm	Beat per minute
BV	Blood volume
CEMA	2-cyanoethylmercapturic acid
CI	Confidence interval
CO	Carbon monoxide
COHb	Carboxyhemoglobin
CRF	Case report form

CRO	Contract research organization
CSP	Clinical Study Protocol
CSR	Clinical study report
CTCAE	Common terminology criteria for adverse events and common toxicity criteria
CV	Coefficient of variation
ECG	Electrocardiogram
EOS	End of study
FAS	Full analysis set
FEV_1	Forced expiratory volume in 1 second
FTND	Fagerström test for nicotine dependence (revised version)
FVC	Forced vital capacity
GH	Growth hormone
Hb	Hemoglobin
HbA1c	Hemoglobin A1c
HDL	High-density lipoprotein
HIV	Human immunodeficiency virus
HR	Heart rate
HS	HeatSticks
hs-CRP	High sensitivity C-reactive protein
ICF	Informed consent form
IP	Investigational product
IPAQ	International Physical Activity Questionnaire
LDL	Low-density lipoprotein
LLOQ	Lower limit of quantification
MedDRA	Medical dictionary for regulatory activities
MHealth	Mobile Health
NEQ	Nicotine equivalents
NNAL	Total 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol

PD	Protocol deviations
PMI	Philip Morris International
PV	Plasma volume
RBCV	Red blood cell volume
RPC	Rating of Perceived Capacity
RPE	Rate of Perceived Exertion
SA	Smoking abstinence
SAE	Serious adverse event
SAP	Statistical analysis plan
SOC	System organ class
SOP	Standard operating procedure
ULOQ	Upper limit of quantification
VAS	Visual analogue scale
VLDL	Very low density lipoprotein
VO_2max	Maximal oxygen uptake
WHO	World Health Organization

3 STUDY OBJECTIVES AND ENDPOINTS

3.1 Objectives and Endpoints

1. To evaluate changes in VO_2max in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)
   Endpoint (V3, V4 and V43)
   - VO_2max as determined during maximal cycle ergometer exercise (expressed in absolute [$mL*min^{-1}$], weight-adjusted [$mL*kg^{-1}*min^{-1}$] and fat free weight-adjusted [$mL*kg^{-1}*min^{-1}$] values)
2. To evaluate changes in exercise capacity in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

Endpoints (V3, V4 and V43)

- Exercise capacity: time to complete a pre-defined work on a cycle ergometer (min:sec)

3. To assess the intensity of the exercise training in subjects participating in a training program and switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

   Endpoints (V5-V42)

   - Cumulative work produced during each training session (calories and calories/kg body weight)
   - Average work rate during each training session (watt and watt/kg body weight)
   - Average work rate during each interval during each training session (watt and watt/kg body weight)
   - Time spent at 0-50%, 50-65%, 65-75%, 75-90%, >90% of maximal work rate during each training session (min:sec)
   - Average heart rate (HR) during each training session (bpm)
   - Time spent at 0-50%, 50-60%, 60-70% , 70-80 % and >80 % of maximal HR during each training session (min:sec)

4. To evaluate changes in physiological parameters and perception of exertion and capacity in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

   Endpoints (V3, V43)

   - Blood composition: Hemoglobin mass (g), red blood cell volume (mL), plasma volume (mL) and total blood volume (mL) as determined by CO re-breathing method
   - Capillary blood lactate levels during VO_2max test (mmol/L)
   - Perceived rate of exertion during VO_2max test (Borg Rating of Perceived Exertion (RPE) scale)

   Endpoints (V3, V4 and V43)

   - Respiratory parameters at VO_2max: Ventilation (L/min), respiratory rate, VCO2 (L/min), respiratory exchange ratio (RER) (VCO2 / VO_2)
   - Rating of Perceived Capacity (RPC) scale
   - Heart rate (bpm) and oxygen uptake i.e. VO_2 (mL/min) during VO_2max test

5. To monitor trends of daily physical activity levels during the study in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

   Endpoints (from V2 until V43)

   Aggregated data derived from the mobile health (mHealth) wearable for:

   - Cumulative number of steps per day
   - Sedentary minutes per day and % of time sedentary

- Active minutes per day and % of time active
- Very active minutes per day and % of time very active

6. To explore trends and changes in activity and sleep parameters through continuous measurements during the study*
   Endpoints (from V2 until V43)
   mHealth wearable measurements of steps, activity, distance, energy expenditure, sleep.
   *Reporting of data for this objective will be subject to a separate report(s) (not part of the Clinical Study Report [CSR]).
7. To describe changes in biological health markers in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)
   Endpoints (V3, V14, V28 and V43)
   - High density lipoprotein (HDL), low density lipoprotein (LDL) and very low density lipoprotein (VLDL) (mg/dL)
   - High sensitivity C-reactive protein (hs-CRP) (mg/L)
   - Growth hormone (GH) (ng/mL)
   - Hemoglobin A1C (HbA1c) (%)
   - Resting blood pressure (mmHg)
   - Resting HR (bpm)
8. To describe changes in weight, body fat and waist circumference in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)
   Endpoints (V3, V4 and V43)
   - Body fat percentage (%)
   - Waist circumference (cm)

   Endpoints (V3- V43)
   - Body weight (kg)
9. To monitor levels of exposure to CO, nitrosamines and acrylonitrile in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)
   Endpoints
   At all visits:
   - Exhaled carbon monoxide (CO) (ppm)

   At V3, V4, V14, V28 and V43
   - Carboxyhemoglobin (COHb%) in blood (%)
   - Nicotine equivalents (NEQ) in urine (adjusted for creatinine)
   - Total 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) in urine (adjusted for creatinine)
   - 2-cyanoethylmercapturic acid (CEMA) in urine (adjusted for creatinine)

10. To describe the self-reported nicotine and tobacco containing product use over the duration of the study
    Endpoints (measured daily from V2 to V43)
    - Self-reported number of any nicotine/tobacco product use on a daily basis as reported in the product use diary
    - Product use exposure
11. To monitor safety during the study
    Endpoints
    - Incidence of adverse events (AEs), serious adverse events (SAEs)
    - Frequency of AEs, SAEs
    - Incidence of IQOS device events including malfunction/misuse
    - Frequency of IQOS device events including malfunction/misuse
    - Vital signs changes from baseline (systolic and diastolic blood pressure, pulse rate and respiratory rate)
    - Physical examination changes from baseline
    - Electrocardiogram (ECG) changes from baseline (heart rate, PR, QRS, QT, QTcB, QTcF intervals)
    - Spirometry changes from screening visit (V1) used as baseline (FEV1, FEV1 % predicted, FVC, FVC % predicted, and FEV1/FVC)
    - Cough assessment changes from baseline (VAS and three Likert scales)
    - Changes from baseline in clinical chemistry, hematology, and urine analysis safety panel
    - Concomitant medications
12. To assess perception of health and functioning in subjects switching to IQOS, continuing to smoke cigarettes and being smoking abstinent **
    Endpoints (first interview between V2 and V3, second interview between V42 and V43)
    - Changes in subjects' perception of health and functioning as determined by qualitative interviews based on the World Health Organization's International Classification of Functioning, Disability, and Health

    ** This objective will be assessed as a sub-study with a separate section of the Informed Consent Form (ICF), and the reporting will be subject of an appendix to the main CSR.

3.2 Exploratory Endpoints

Not Applicable.

3.3 Study Hypotheses and Evaluation Criteria

3.3.1 Hypotheses

This study is descriptive and therefore no statistical hypothesis will be tested.

4 INVESTIGATIONAL PLAN

4.1 Study Design

This is a randomized, controlled, open-label, 4-arm parallel group study with stratification by daily cigarette consumption over the past 12 months prior to V1 as reported during V1 (i.e. 10 to 19 cigarettes or >19 cigarettes per day) and sex (Figure 1).

A sufficient number of healthy adult smokers will be screened and enrolled after checking that all eligibility criteria have been met, in order to reach 90 randomized subjects.

Smokers will be randomized as follows with a 5:5:5:3 ratio according to strata:

1. switch to IQOS use + participation in training program: IQOS-1 arm, 25 subjects
2. continue cigarette smoking + participation in training program: Cigarette arm, 25 subjects
3. smoking abstinence + participation in training program: SA arm, 25 subjects
4. switch to IQOS use only: IQOS-2 arm, 15 subjects

When 90 randomized subjects are reached, further enrollment will be stopped. Subjects already enrolled in the study will still be eligible for randomization. From randomization, subjects will be instructed to use their allocated product or stay smoking abstinent until the end of the study. Drop-outs will not be replaced.

Acute effect on VO2max and exercise capacity of switching to IQOS will be assessed one week after randomization before initiating the training program, and the combined effect of switching and exercise training will be assessed one week after completion of the 12-week training program.

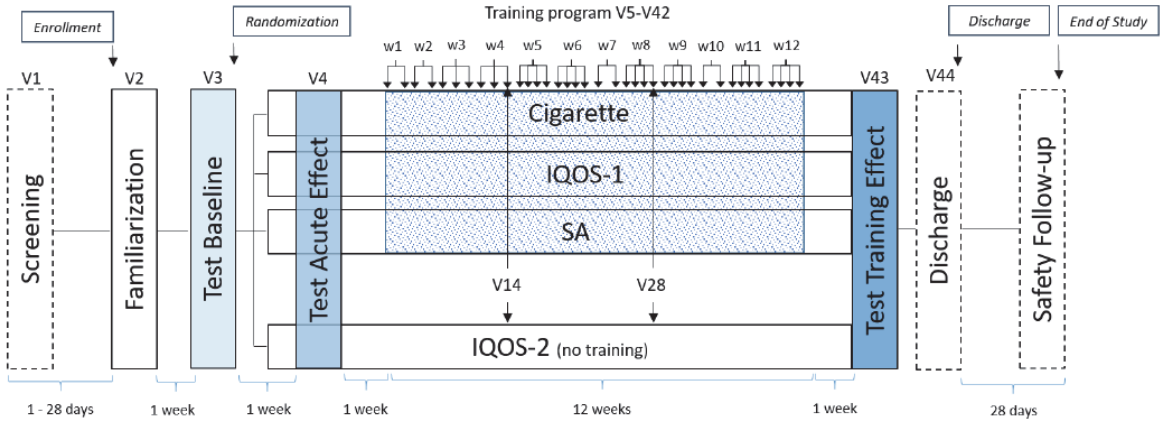

Figure 1 Study Design

Subjects will have serial visits at the investigational site as follows:

Screening, V1 (within 28 days prior to V2): No study-specific procedures will be performed before the Informed Consent Form (ICF) has been signed. The subject will also be invited to participate in a sub-study on perception of health and functioning and will sign a separate optional section of the ICF for this part of the study. Subjects will be invited to sign a separate optional consent form for collection of samples for biobanking.
A demonstration of *IQOS* and the mHealth wearable will be done by the site staff during the screening visit. All inclusion and exclusion criteria will be assessed. A VO_2max test will be performed as part of eligibility assessments. All eligible subjects will be invited to come to V2.

Enrollment and Familiarization, V2 (7±1 days prior to V3): Enrollment will take place after confirmation of smoking status (urine cotinine and exhaled CO) and negative urine pregnancy test, alcohol breath test and urine drug screen. After enrollment, vital signs will be assessed and subjects will perform an exercise capacity test for the purpose of familiarization. The exercise test consists of completing a pre-defined work (calories) on a bike ergometer (1). Before leaving the investigational site the subject will receive a mHealth wearable kit and will be trained on how to handle the device. Subjects will be instructed to wear the device all the time. Subjects will also receive a diary to record the daily use of tobacco and nicotine containing products.
Between V2 and V3, subjects having signed the optional section of the ICF will have an interview (by phone or computer) on the topic of health and functioning lasting around 60 minutes.

Baseline and Randomization, V3: Baseline assessments will be performed as listed in Appendix 1. Then, 30-60 min before start of the VO_2max test all subjects will smoke a cigarette (only if they are willing to). During the VO_2max test at rest and at each 25 W increase, the subject's lactate levels will be determined by capillary blood sampling and the rate of perceived exertion (RPE) will be assessed using the Borg RPE scale. After completing the VO_2max test, the subject will rest for 60±5min, after which the exercise capacity test will be performed. After completing exercise tests, total blood volume will be determined.
Subjects will thereafter be randomized into one of the 4 arms. Subjects randomized into *IQOS* arms will receive their *IQOS* starter kit and be trained on how to use the device. Subjects randomized into the SA arm will receive smoking abstinence support and will be allowed to use NRT upon request.

Acute effect, V4 (7±2 days after V3): Before the VO_2 max test, vital signs, exhaled CO, weight, body-fat and waist circumference measurements will be performed. 30-60 min before start of the VO_2max test subjects in cigarette and *IQOS* arms will smoke a cigarette or use a HeatStick, respectively (only if they are willing to). Subjects' VO_2max, maximal heart rate and maximal work rate will be determined. After completing the VO_2max test, subjects will rest for 60±5min after which the exercise capacity test will be performed.

Training program, from V5 (15±2 days after V3) until V42 (99±2 days after V3): Subjects randomized into *IQOS*-1, cigarette and SA arms will participate in 38 supervised training sessions during a 12-week period using three different training protocols (Appendix 2).

Subjects randomized to the *IQOS*-2 arm will come for visits once a week for all assessments except the training. Subjects in the *IQOS*-2 arm will consequently only attend the following visits: V5, V7, V9, V14, V16, V19, V23, V28, V30, V33, V35 and V39.

At V14 and V28 (medical visits), safety, biological health marker and exposure marker assessments will be performed for all subjects.

Before each training, the following assessments will be conducted;

- Weight
- Exhaled CO
- 30-60 min before start of the training subjects in cigarette and IQOS arms will smoke a cigarette or use a HeatStick, respectively (only if they are willing to).

All trainings will be performed on a bike ergometer. The training intensity will be adjusted according to the subject's resting heart rate as determined by ECG at V3, and maximal heart rate as determined at the VO_2max test at V4 (acute test) . The work produced, the work rate and heart rate will be recorded for each training (as described in Objective 3).

Between V42 and V43, subjects having signed the optional section of the ICF will have an interview (by phone or computer) on the topic of health and functioning lasting around 60 minutes.

Training effect, V43 (106±2 days after V3):
Before the VO_2max test, biological health markers, exposure markers, vital signs, exhaled CO, weight, body-fat percentage and waist circumference measurements will be performed. Then, 30-60 min before start of the VO_2max test subjects in cigarette and *IQOS* arms will smoke a cigarette or use a HeatStick, respectively (only if they are willing to). Subjects' VO_2max will be determined as described for V3, including determination of lactate levels and perceived exertion rate. After completing the VO_2max test, the subject will rest for 60±5min, after which the exercise capacity test will be performed.

After completing exercise tests, total blood volume will be determined. Subjects will return the mHealth wearable, the product use diary and the *IQOS* device (if applicable).

Discharge, V44 (107±2 days after V3): Safety assessments will be conducted as listed in Appendix 1 of the study protocol.

Safety follow-up period (from discharge at V44 or early termination plus 28 days):
After discharge on V44 or after early termination, subjects will enter a 28-day safety follow-up period during which AE/SAEs spontaneously reported by the subjects will be collected. Any non-serious AE that is ongoing at the time of discharge or early discontinuation will be followed-up by the Investigator or designee during the safety follow-up period until it has been resolved, stabilized (i.e., no worsening of the condition), or an acceptable explanation has been found (e.g., a chronic condition) or lost to follow up. The follow-up of the ongoing non-serious AEs will be done via a phone call performed until the end of the Safety Follow-Up Period. At the end of the safety follow-up period, all ongoing non-serious AEs will be documented as "ongoing" and no further follow-up information will be sought for them by the Investigator or designee. At that point, the Investigator will assess whether the subject should be referred to his/her General Practitioner to have their ongoing AEs addressed accordingly. SAEs will be followed up by the Investigator until resolution, stabilization or the determination of a plausible explanation for them was found, regardless of the end of the safety follow-up period. All subjects discontinued from the study at any time after enrollment, will enter the 28-day safety follow-up period.

4.2 Selection of Study Population

4.2.1 Inclusion Criteria

Subjects who meet all the following inclusion criteria can be enrolled into the study:

Inclusion Criteria	Rationale	Screening	Enrollment V2
1. Subject has signed the ICF.	Administrative	x	
2. Smoking, healthy subject based on safety laboratory, ECG, spirometry, vital signs, physical examination, medical history and Investigator's assessment.	Safety	x	
3. Subject has been smoking for at least three years prior to V1.	Effect	x	
4. Subject has been smoking ≥ 10 cigarettes per day over the last 12 months. Smoking status will be verified by a urinary cotinine ≥ 200 ng/mL and CO exhaled breath test > 6ppm both at V1 and V2.	Effect	x	x
5. Subject does not plan to quit smoking within 6 months after V1.	Safety	x	

Inclusion Criteria	Rationale	Screening	Enrollment V2
6. Subject is aged between 21 and 65 years (inclusive).	Safety	x	
7. Subject is available for the entire study period and willing to comply with study procedures.	Effect	x	

4.2.2 Exclusion Criteria

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

Exclusion Criteria	Rationale	Screening	Enrollment V2
1. Subject has a clinically relevant disease which requires medication (including but not limited to gastrointestinal, renal, hepatic, neurological, hematological, endocrine, oncological, urological, immunological, pulmonary [such as but not limited to pulmonary oedema, asthma], and cardiovascular [such as, but not limited to myocardial infarction, unstable angina, uncontrolled arrhythmias, heart failure], disease) or any other clinically significant medical condition (including abnormal safety laboratory result as per CTCAE), which as per the judgment of the Investigator would jeopardize the safety of the subject.	Safety	x	
2. Subject who has forced expiratory volume in 1 second/forced vital capacity (FEV_1/FVC) <0.7 and FEV_1 <80% predicted value at post-bronchodilator spirometry (GOLD, 2017).	Safety	x	
3. Subject with asthma condition (FEV_1/FVC < 0.75 and reversibility in FEV_1 (in both > 12% and > 200 mL) from pre to post-bronchodilator values.	Safety	x	
4. Subject has clinical significant abnormalities of ECG at V1.	Safety	x	
5. Subject performs more than 45 min of vigorous physical activity per week.	Effect	x	
6. Inability to perform a VO_2max test at V1.	Safety	x	
7. VO_2max >50 $mL.min^{-1}kg^{-1}$ for men and VO_2max >40 $mL.min^{-1}kg^{-1}$ for women as determined at V1.	Effect	x	
8. Subject takes medication influencing blood volume such as erythropoietin, diuretics and beta blockers, or diabetic medications.	Effect	x	

Exclusion Criteria	Rationale	Screening	Enrollment V2
9. Subject cannot participate in the study for any reason other than medical as per the Investigator's judgment (e.g. psychological and/or social reason)	Administrative	x	
10. For women only: subject is pregnant (does not have negative pregnancy tests at V1 and at V2) or is breastfeeding.	Safety	x	x
11. For women of childbearing potential[1]: female subject who does not agree to using an acceptable method of effective contraception[2] during the entire study.	Safety	x	
12. Subject has a BMI < 18.5 kg/m^2 or BMI ≥ 30 kg/m^2.	Safety	x	
13. Subject has positive serology test for HIV, Hepatitis B or Hepatitis C.	Safety	x	
14. Subject has a positive alcohol breath test and/or a history of alcohol use disorder (both V1 and V2).	Administrative	x	x
15. Subject has a positive urine drug screen.	Administrative	x	x
16. The subject has been previously screened for this study.	Administrative	x	
17. The subject, or one of their family members (e.g., spouse, parent, sibling or child), is a current or former employee of the tobacco industry.	Administrative	x	
18. The subject, or one of their family members (e.g. spouse, parent, sibling or child), is an employee of the investigational site or any other parties involved in the study.	Administrative	x	
19. Subject is legally incompetent, or physically or mentally incapable of giving consent (e.g., emergency situation, under guardianship, prisoners, or subjects who are involuntarily incarcerated).	Administrative	x	

[1] Women who are not of childbearing potential meet at least one of the following criteria: Have undergone hysterectomy or bilateral tubal ligation, Have medically confirmed ovarian failure, or Are medically confirmed to be post-menopausal (cessation of regular menses for at least 12 consecutive months with no alternative pathological or physiological cause).

[2] Intrauterine device, intrauterine system, barrier methods of contraception (condoms, occlusive caps) with spermicidal foam/gel/film/suppository, hormonal contraception containing progesterone only, vasectomized partner(s), from screening until the end of the safety follow-up period.

4.3 Product Allocation and Blinding

4.3.1 Method of Assigning Subjects to Sequence/Product Arms

Randomization will be done through the [redacted] at any time during the baseline visit. Subjects will be informed about their study arm allocation at the end of V3. Subjects will be randomized in one of the four study arms:

- Cigarette arm with training program
- IQOS-*1* arm with training program
- SA arm with training program
- IQOS-*2* arm without training

in a 5:5:5:3 ratio using a stratified randomization based on daily cigarette consumption over the 12 months prior to V1 (10-19 cigarettes/day vs. >19 cigarettes/day), and sex. All subjects will be instructed to adhere to their assigned product/regimen until they complete the study.

4.3.2 Blinding

4.3.2.1 Blinding of site staff

This is an open-label study and the site staff performing study assessments will therefore not be blinded to the randomized arm of the subject.

4.3.2.2 Blinding of data

There will be additional, even though limited, degree of blinding during the conduct of the study, including the data review and data analysis process prior to database lock. In particular, PMI and contract research organization (CRO) personnel will be blinded as summarized in Table 1.

Table 1: Blinding

Blinded Study Personnel	Blinded data	End of Blinding Period
PMI and CRO Study Statisticians	VO_2max and exercise capacity test results after randomization	After database lock.
PMI Clinical Scientist	VO_2max and exercise capacity test results after randomization	After the finalization of PMI blind database review. Can be actively un-blinded when appropriate.

Any PMI and CRO personnel who are not listed in Table 1 will by default be not be blinded.

Unblinded information will not be shared with the blinded study team, until the end of the blinding period. PMI will receive blinded and unblinded data for the pre-analysis data review as planned in the data review plan. Unblinded data will only be reviewed by the unblinded study team.

4.3.3 Compliance to Product Allocation

Subjects will enter the number of HeatSticks, cigarettes, or/and other tobacco and nicotine containing products used per day in a product use diary from V2 until the discharge at V43.

5 DERIVED AND COMPUTED VARIABLES

5.1 Urinary Biomarkers

Spot urine will be collected to measure concentrations of NEQ, NNAL, CEMA and creatinine.

5.1.1 Nicotine Equivalents

The concentration of NEQ in spot urine will be derived according to the formula below, and considering the conversion factors described in Table 2: Conversion factors from ng/ml into µmol/L. The concentrations reported for free nicotine and its five major metabolites will not be used individually as analysis variables.

$$\begin{aligned}\text{Neq [mg/L]} = {} & (\text{free nicotine[µmol/L]} + \text{nicotine-glucuronide [µmol/L]} \\ & + \text{free cotinine [µmol/L]} + \text{cotinine-glucuronide [µmol/L]} \\ & + \text{free } \textit{trans}\text{-3'-hydroxycotinine [µmol/L]} \\ & + \textit{trans}\text{-3'-hydroxycotinine-glucuronide [µmol/L])}\end{aligned}$$

*162.2[µg/µmol] / 1000

Table 2: Conversion factors from ng/ml into µmol/L

	Molecular weight (g/mol)	**Conversion factor from ng/mL to µmol/L**
Free Nicotine	162.232	0.006164
Nicotine glucuronide	338.356	0.002955
Free Cotinine	176.218	0.005675
Cotinine-glucuronide	352.341	0.002838
Free *Trans*-3'hydroxycotinine	192.217	0.005202
Trans-3'hydroxycotinine-glucuronide	368.34	0.002715

5.1.2 Biomarkers of Exposure adjusted for Creatinine

The adjustment for creatinine for the urinary BoExp will be calculated as:

$$\left[\text{BoExp (creatinine adjusted)} = \frac{\text{BoExp}}{\text{Creatinine}}\right]$$

NEQ, Total NNAL, CEMA will be expressed in mg/g creat, pg/mg creat and ng/mg creat respectively.

5.2 Blood composition as determined by CO re-breathing method

Blood composition variables as determined by CO re-breathing method will be calculated according to the method proposed in Siebenmann (Siebenmann, et al. 2017). The corresponding endpoints are hemoglobin mass (g), red blood cell volume (mL), plasma volume (mL) and total blood volume (mL).

The change in the fraction of carboxyhemoglobin in blood samples obtained pre- and post-CO rebreathing is calculated in percentages as:

ΔCOHb = %COHbPost - %COHbPre.

The absorbed number of CO molecules is calculated in moles as:
nCOabsorbed = Patm * VCOabsorbed/(R * T)
where R is the ideal gas constant [0.008206 L*atm/(mol*K)]. The volume of CO that has been absorbed (VCOabsorbed), atmospheric pressure (Patm), and room temperature (T) should be entered in liters, atmospheres, and Kelvin, respectively.

The value of the volume of CO that is absorbed is normalized to 25 C° and 1013 mbar (i.e. 1 atm) and called VCOabsorbed, norm. Thus in the equation above Patm = 1 atm and T = 298.15 K. The absorbed number of CO molecules is therefore,

nCOabsorbed = 1* VCOabsorbed,norm / (82.06 *298.15) = VCOabsorbed,norm / 24466.189.

Given that 1 mol of hemoglobin binds 4 mol of CO the tagged number of hemoglobin molecules (nHbtagged) is calculated in moles as:

nHbtagged = nCOabsorbed/4.

The total number of hemoglobin molecules is calculated in moles as:
nHbtotal = (nHbtagged/ΔCOHb) * 100.

Based on the molar mass of hemoglobin, the hemoglobin mass (in grams) is calculated as:
Hbmass = nHbtotal * 6.44 * 10^4 g/mol.
With the equations above, a VCOabsorbed,norm of 90 mL and a ΔCOHb of 10 % results in 592.25 g HBmass.

Intravascular volumes (in mL) are calculated as:
Red blood cell volume, i.e. RBCV = Hbmass * Hct/[Hb]
Total blood volume, i.e BV = RBCV/Hct
Plasma volume, i.e PV = BV - RBCV
Hct is entered in these equations as absolute fractions (i.e., not in %), whereas [Hb] is entered in g/mL.

5.3 Questionnaires

5.3.1 Fagerström Test for Nicotine Dependence (Revised Version)

Subjects will complete the FTND questionnaire in its revised version (Heatherton, et al. 1991).

The questionnaire consists of six questions. The FTND total score determines the subject's dependence on nicotine in three levels: mild (0-3 points), moderate (4-6 points), and severe (7-10 points).

The responses are scored as defined in the Table 3 below and the sum score (FTND score) will be calculated.

Table 3: Scoring of Responses to the Fagerstöm Test for Nicotine Dependence

Item No.	Question	Response	Score
1	How soon after you wake up do you smoke your first cigarette?	within 5 min	3
		6-30 min	2
		31-60 min	1
		after 60 min	0
2	Do you find it difficult to refrain from smoking in places where it is forbidden?	Yes	1
		No	0
3	Which cigarette would you hate most to give up?	The first one in the morning	1
		Any other	0
4	How many cigarettes do you smoke per day?	1 to 10	0
		11 to 20	1
		21 to 30	2
		31 to 40	3
		More than 40	3
5	Do you smoke more frequently during the first hours after waking than during the rest of the day?	Yes	1
		No	0
6	Do you smoke if you are so ill that you are in bed most of the day?	Yes	1
		No	0

5.3.2 Product Preference

On V3, the following question will be asked to all subjects:

"Which product would you prefer to be randomized to?"

- IQOS
- Cigarette
- Smoking abstinence
- No preference

5.3.3 Lifestyle questionnaire

Subjects will be asked questions to capture baseline covariates such as, diet, alcohol intake, sleep deficit, and exposure to passive smoking at V3.

5.3.4 Assessment of Cough-VAS and Likert Scales

Subjects will assess cough on a Visual Analogue Scale (VAS) and on three Likert scales.

Subjects will be asked if they have experienced a regular need to cough (*e.g.*, 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 a VAS and 3 Likert scales. On the VAS, subjects will assess how bothersome their cough was during the previous 24 hours. The VAS ranges from "not bothering me at all" to "extremely bothersome". Furthermore, subjects will assess the intensity and frequency of cough and the amount of sputum production during the previous 24 hours on Likert scales.

- The intensity of cough will be assessed on a 5-point Likert scale ranging from 1 to 5, with 1 = very mild – 2 = mild – 3 = moderate – 4 = severe – 5 = very severe.
- The frequency of cough will be assessed on a 5-point Likert scale ranging from 1 to 5, with 1 = rarely – 2 = sometimes – 3 = fairly often – 4 = often – 5 = almost always.
- The amount of sputum production will be assessed on a 4-point Likert scale ranging from 0 to 3, with 0 = no sputum – 1 = a moderate amount of sputum – 2 = a larger amount of sputum – 3 = a very large amount of sputum.

5.3.5 International Physical Activity Questionnaire (IPAQ) - short form

At V1 subjects' physical activity levels will be assessed using the short version of the IPAQ. The short form records the activity of four intensity levels:

- vigorous-intensity activity such as aerobics,
- moderate-intensity activity such as leisure cycling,
- walking, and
- sitting.

The categorical score of the short question of the IPAQ will be presented using the following algorithm, further details can be found in (IPAQ 2005).

1. Low

No activity is reported OR
Some activity is reported but not enough to meet Categories 2 or 3.

2. Moderate

Either of the following 3 criteria
• 3 or more days of vigorous activity of at least 20 minutes per day OR
• 5 or more days of moderate-intensity activity and/or walking of at least 30 minutes per day OR
• 5 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 600 MET-minutes/week.

3. High
Any one of the following 2 criteria
• Vigorous-intensity activity on at least 3 days and accumulating at least 1500 MET-minutes/week OR
• 7 or more days of any combination of walking, moderate- or vigorous-intensity activities accumulating at least 3000 MET-minutes/week

5.3.6 Borg Rate of Perceived Exertion (RPE) Scale

During VO_2max test (only at V3 and V43) subjects will rate their perceived exertion on the Borg RPE scale at rest and then at each 25W increase (Borg 1982). The Borg RPE scale is a way of measuring physical activity intensity level. It is based on the physical sensations a person experiences during physical activity, including increased heart rate, increased respiration or breathing rate, increased sweating, and muscle fatigue.

The scale ranges from
- 6- no exertion at all,
- 7.5 – extremely light,
- 9 – very light,
- 11 – light,
- 13 somewhat hard,
- 15- hard,
- 17-very hard,
- 19- extremely hard, to
- 20- maximal exertion.

A high correlation exists between a person's perceived exertion rating times 10 and the actual heart rate during physical activity.

5.3.7 Rate of Perceived Capacity (RPC) Scale

To predict maximal exercise capacity, the rating of perceived capacity (RPC) scale was developed based on metabolic equivalents (METs) (Wisen, Farazdaghi and Wohlfart 2002).

MET values from 1 to 20 for men and 1 to 18 for women are listed on a progressive scale and linked to physical activities. Subjects will rate their perceived capacity by choosing the most strenuous activity and the corresponding MET value that they could sustain for 30 min.

5.3.8 Product Use Diary

Subjects will enter the number of HeatSticks, cigarettes, or/and other tobacco and nicotine-containing products used per day in a product use diary from V2 until discharge at V43.

5.3.9 Qualitative Interviews on Health and Functioning (sub-study)

Subjects that have signed the optional section of the ICF for the sub-study on health on functioning will have a one-hour, qualitative interview over the phone or computer between V2 and V3, and at the end of the study between V42 and V43. Questions during the interview will be open-ended and will focus on subjects' current state of health and functioning, experiences during the training program, and changes in health and functioning as a result of the training program.

5.4 Categorical Variables

For categorical variables used in this study see Section 9.1.1 "Stratified Presentation".

5.5 Product Use

The number of Cigarettes, Heatsticks, and other nicotine/tobacco containing product used daily, as reported by the subject and collected in the electronic diary or paper CRF, will be used to monitor product use and evaluate adherence to product allocation.
Additionally, BoExp measured in the study will be used to fine tune product use allocation as an exploratory analysis.

5.5.1 Product Use Pattern

Actual product use pattern categorization is described in Figure 2 below,

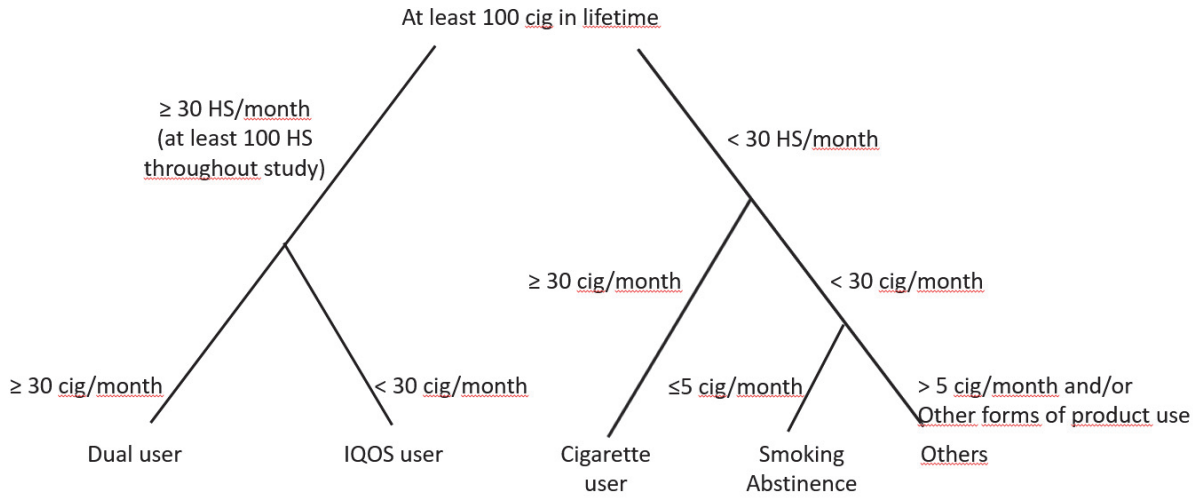

Figure 2: Product Use Categorization

Note: HS: HeatSticks

The calculation of product use categories will be performed based on Figure 2. For each subject we will calculate product use for two time periods, representing Week 1 use and average product consumption from V3 till end of Month 3 or last visit prior to early termination whichever comes first.

- From V3 to V4: Week 1
- From V3 to V43/Last visit before early termination: Overall average consumption

For each time period the daily product use reported by a subject in the study will be summed up and averaged by the actual number of days the subject was in the study after randomization until date of last reported product use within the respective time period. This will then be multiplied by 30 to describe the average monthly product use.

For subjects who early terminated from the study, product use for overall average consumption will be calculated until the last visit prior to V44. For example, if a subject early terminated at V23 and subsequently came for V44, then product use will be calculated from V3 until the date of V23.

As both electronic diaries and paper diaries were used in the study concurrently for some subjects, the following rules will be applied in case of inconsistencies.

- When there is data available for both electronic and paper diaries for the same date, the electronic diary data will be kept if both are the same. If the data is different on the same date, then data for this date will be considered missing.
- If a subject reports no product use but a value is recorded, we will assume that the number of product used reported was correct and use it in the derivation of product use.

Missing data for self-reported daily use will not be imputed and we will assume that the derived average monthly use from non-missing data is representative of the actual usage. For each analysis period (Week 1 or Overall), if more than 50% of product use data is not available then the subject will be grouped into the 'Others' product use arm. Therefore, it is possible that a subject is not considered for statistical modelling as it is grouped into 'Others' for Week 1 (3 out of 7 days of product use data only) but available for analysis for the entire study period if more than 50% of product use data was eventually collected throughout the study. A subject may have switched product use groups throughout the study and the total number of subjects in an arm may fluctuate over time depending on the actual product use.

5.5.2 Cigarette Use

The average number of cigarettes used by each subject from familiarization V2 to V3 will be calculated, in addition the average from baseline (V3) to the following visits will be calculated, these visits include V4, V14 (Month 1), V28 (Month 2), V43 (Month 3) which corresponds to cigarettes smoked from baseline to acute effect visit, end of Month 1, end of Month 2 and end of Month 3 respectively.

5.5.3 IQOS Use

The average number of IQOS Heat sticks used by each subject from familiarization V2 to V3 will be calculated, in addition the average IQOS use from baseline (V3) to the following visits will be calculated, these visits include V4, V14 (Month 1), V28 (Month 2), V43 (Month 3) which corresponds to Heat sticks used from baseline to acute effect visit, end of Month 1, end of Month 2 and end of Month 3 respectively.

6 SAMPLE SIZE JUSTIFICATION

The sample size of 20 per arm (smoking abstinence with training, IQOS with training, continued cigarette smoking with training) was determined based on previously published results, which reported a standard deviation of 14 mL/kg/min in VO_2max in the change

between groups which continued smoking and those who were smoking abstinent after participating in a 12-week training program (Albrecht, et al. 1998). This sample size allows a probability of 90% that the half-width of the 95% confidence intervals between two arms have a precision of approximately 10.3 mL/kg/min.

For example, if the mean difference in change between the cigarette arm and smoking abstinence arm was 12 mL/kg/min after training, we expect that if the experiment were repeated many times then on average 90% of the time the 95% confidence intervals of the mean differences between two arms will be within 1.7 and 22.3. To allow for a dropout rate of approximately 20%, the sample size will be 25 per arm. The IQOS-*2* arm without training will mainly serve as a secondary analysis to the IQOS-*1* arm with training, and the drop-out rate in this arm is expected to be lower. Therefore, only 15 subjects will be randomized to this arm.

7 CHANGES FROM THE PROTOCOL SPECIFIED STATISTICAL ANALYSES

An exploratory analysis of VO2max and exercise capacity will be carried out on subjects with additional biochemical verification of smoking status.

8 ANALYSIS SETS

The main population for non-safety analysis will be the As Exposed Exercise Compliant Set. Safety will be analyzed using the Safety Set.

8.1 Full Analysis Set

The Full Analysis Set (FAS) consists of all randomized subjects who have at least one post-randomization product (cigarette or IQOS) use experience if randomized to IQOS, IQOS-*2* or cigarette arm, and who have at least one valid non-safety assessment after randomization. All subjects in the SA arm who fulfil these requirements, except the product use experience, are part of this set as well. The FAS will be analyzed by randomized study arm.

8.2 Safety Set

The Safety Set consists of all subjects enrolled with signed ICF who have at least one valid safety assessment during the course of the study. The Safety Set will be analyzed by actual exposure (product use exposure, actual study arm).

8.3 As Exposed Exercise Compliant Set

The As Exposed Exercise Compliant Set (AEECS) consists of all randomized subjects who have at least one valid non-safety assessment after randomization, specifically we have

- post-randomization product (cigarette or IQOS) use experience if randomized to IQOS, IQOS-2 or cigarette arm or
- no product use if randomized to the SA arm,

and who attended at least 34 of 38 (~90%) of the exercise training sessions except the subjects in IQOS-*2* arm. Subjects in the IQOS-*2* will be analyzed as exposed as they did not participate in the training. The AEECS will be analyzed by actual exposure (product use exposure, actual study arm).

Due to technical issues with the bike training equipment during the study, these rules are used to determine if a subject completed a training session for the purposes of defining the AEECS analysis set in addition to information regarding whether a training visit was done or not. We define a training session to be completed if a minimum of 20 or more minutes of exercise was done. If the training time is unknown then the session is considered to be not done. These checks from the protocol deviation dataset will be done regardless of whether data from the exercise training is available or not.

For a training session to be considered as not done from the protocol deviations (PD) dataset, the PD description will be checked for the following,

- “Bike training was not done …”
- “Bike training was partially completed for less than 20 min …”
- “Bike training was partially completed for unknown number of minutes …”

A bike training session is considered as complete if the following description for bike training is provided in the PD dataset,

- “Bike training was completed…”
- “Bike training was partially completed for at least 20 min…”

In the case where there are inconsistencies between the PD dataset and bike training data, ie, for example bike training was marked as complete but the PD descriptions says otherwise, the PD rules will be used to determine completion of a bike training session.

8.4 As Exposed Biochemically and Exercise Compliant Set

A subset of the AEECS based on biochemical verification of smoking status (AEBECS) will be analysed based on actual arm as determined with these additional criteria.

For a subject to be classified as smoking abstinent, in addition to the product use diary with an average cigarette use at max 5 Cigarettes/month the following will be used,

- No CO breath test > 6 ppm at three consecutive visits from V3 onwards.
- Total NNAL concentration < 75.9 pg/mL in spot urine **(Berg, et al. 2012)** at two or more visits collected from V3 onwards.
- CEMA concentration adjusted for creatinine < 40 ng/mg creat **(Claussen, et al. 2019)** in spot urine at two or more visits collected from V3 onwards.

The following biochemical verification will be used in addition to the product use diary to further fine tune subjects in the IQOS user group.

- No CO breath test > 6 ppm at three consecutive visits from V3 onwards.
- Total NNAL concentration < 75.9 pg/mL in spot urine at two or more visits collected from V3 onwards.
- CEMA concentration adjust for creatinine < 40 ng/mg creat in spot urine at two or more visits collected from V3 onwards.

8.5 Protocol Deviations

All protocol deviations will be listed and also summarized based on the classification of deviation type for subjects in the Safety Set.

8.5.1 Major Protocol Deviations

Major protocol deviations are protocol deviations that may significantly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety, or well-being.

Subjects with major PDs below will be identified to determine whether they will be excluded from any of the analysis sets.

The following general rules will be applied to determine if a deviation was major:

1. Protocol violations pertaining to inclusion/exclusion criteria will considered major deviations, evaluable if no results or safety is impacted.
2. Protocol violations pertaining to ICF will be considered major deviations, evaluable if no results or safety is impacted. If the ICF was not signed then the subject will be major, non-evaluable.
3. If a subject was mis-randomized, defined as being administered to the wrong product according to the randomization schedule during the exposure period.
4. The subject did not attend a scheduled visit.
5. The subject did not record more than 50% of product use questionnaires for an analysis period.

8.5.2 Minor Protocol Deviations

Minor protocol deviations are defined as deviation that does not impact any of the following: (1) the safety of a subject participating the study; (2) the integrity of the data collected for the evaluation of the study; or (3) the overall evaluation and/or interpretation of the outcome of the study, specifically with respect to the primary objective(s).

9 PLANNED STATISTICAL METHODS

9.1 General Considerations

The statistical evaluation will be performed using SAS®, version 9.3 or later.

Data presentation formats, precision and rounding will consider Section 2.5 of Record ID PMI-RRP-FOR-112420. Continuous variables will be presented with 3 significant digits, however, values above 999 given as integers and will not be rounded to have zeroes at the rightmost digit(s). For example 12345.67 will be presented as 12346 and not as 12300.

9.1.1 Stratified Presentation

Stratification criteria for the evaluation are:

- Randomized study arm Cigarette, IQOS-*1*, smoking abstinence (SA), IQOS-*2*) - randomized arm.
- Product use (actual arm) i.e. actual product use exposure categorized based on self-reporting and biochemical verification according to Section 5.5.1 “Product Use Pattern”:

- Cigarette,
- IQOS-*1*,
- Smoking Abstinence (SA),
- IQOS-*2*,
- Dual users considered for IQOS-*1* (Dual-1),
- Dual users considered for IQOS-*2* (Dual-2),

- Cigarette consumption at V1 (10 to 19 cigarettes, more than 19 cigarettes per day)
- Sex (male, female)
- Age (<50, ≥ 50 years i.e. 21≤Age<50 and 50≤Age≤65])

The main stratification criterion for the evaluation is the actual product exposure, whereas data will be listed by randomization arm.
Further stratification by cigarette consumption and sex will be done for the analysis of VO_2max and exercise capacity and where explicitly defined. Further stratification by age and BMI will be done only where explicitly defined (e.g. for demographic data).
The AEECS, AEBECS and the Safety Set will be analyzed by actual study arm (product use exposure). The FAS will be analyzed by randomized study arm.

9.1.2 Descriptive Statistics

Descriptive statistics for continuous variables will include the number of subjects, number and percent of subjects with missing data, the mean and standard deviation, geometric means and coefficient of variation (CV), median, first and third quartiles (Q1 and Q3), minimum and maximum, and 95% confidence interval (CI; based on a t-distribution if not otherwise stated) on the mean for each study arm, and summary across all subjects if appropriate. For categorical variables descriptive statistics will provide the number of observations, absolute frequencies, and relative frequencies (percentages) of the observed values. In addition, the results may be presented as a stratified summary.

9.1.3 Definitions for Statistical Data Analysis

The baseline value for any given parameter will be the last assessment prior to randomization for the subjects randomized to either IQOS, cigarette or SA arms.

Change from baseline (i.e. difference of post baseline value compared to baseline value) will be calculated as: Change from baseline = post baseline value – baseline value.

For log-transformed variables, the (relative) change from baseline is defined as: Change from baseline = log(post baseline value / baseline value).

For step count, the baseline will be the average steps taken between V2 and V3.

Time variables, i.e. exercise capacity and time spent at different percentages of maximal work rate (or maximal HR) during each training session will be analyzed in seconds (not in min:sec).

9.1.3.1 Analysis Period

In general, the analysis will be performed at a visit(s) level but in the case where data is collected more frequently for example the Mhealth or exercise training data we define the analysis period by month. Here, Week 1 is the period between V3 and V4, Month 1 is described as the period between V5 up to the date of V14, Month 2 is the day after V14 and until the date of V28 and Month 3 represents the day after V28 until the date of V43.

For certain endpoints (see section 9.6.1.2) related to $VO2_{max}$ and exercise capacity the analysis period Week 1 and Month 3 will have different baseline values dependent on the subject's product use group during the corresponding period for the calculation of the descriptive statistics.

9.1.4 Handling of Dropouts, Missing Data (Including Outside the Limits of Detection or Quantification) or Partial Data

Dropouts (discontinued subjects) and other subjects with missing data will be analyzed in the analysis set(s) when data are available, without imputation of missing data.

No imputation schemes for missing values will be applied.

For laboratory parameters outside the limit of detection or quantification, the following imputation will be performed:

- Values below the lower limit of detection (LLOD) or quantification (LLOQ) will be imputed as 0.5 x LLOD or LLOQ.
- Values above the upper limit of detection (ULOD) or quantification (ULOQ) will be imputed as the ULOD or ULOQ.

The number of values below LLOD (or LLOQ) or above ULOD (or ULOQ) will be presented in each summary table, as well as the statistics on the other quantitative values.

However, if 50% or more data are below the lower limit or above the upper limit, only the number of values below the lower limit and above the upper limit will be reported in the

summaries, together with minimum (if no value below the lower limit is present) and maximum (if no value above the upper limit is present) of the observed values, and no other statistics will be reported.

There will be no imputation of missing laboratory parameters.

For product use data that are missing:

- There will be no imputation and all available data (ePRO and/or paper) will be analyzed as is.

For bike training data that is incomplete:

A bike training session is defined as being complete if a subject trained for at least 20 minutes, if the training stop time is not available then the training session is taken as not done.

For heart rate data that was incomplete (<80 %) during V5, V14, V29 and V41 will be flagged and not used in the analysis of % of time spend at maximum heart rate. This will be flagged by using the protocol deviation description “Recording of HR is less than 32 min …”.

9.1.5 Handling of Unplanned Data

The following rule will be applied to handle unscheduled visits: In the case of unscheduled measurements prior to randomization, the last value before randomization will be considered as baseline value. In all other cases, the unscheduled value will not be used for descriptive statistics as long as a corresponding scheduled value is available. Unscheduled measurements will be included (and marked) in the listings.

9.1.6 Handling of Mhealth Data

For Mhealth parameters, all data will be analyzed as is and no imputation will be performed. However, the first and last day will not be used in any analysis as the data may not be complete.

In addition, the following definitions will be used to define a valid day in terms of step count. Using the average of 5205 steps per day taken in Germany (Althoff, et al. 2017), we set a cut-off at approximately 20% and define a valid day as having more than or equal to 1000 steps. In a separate study on Multiple Sclerosis (MS) patients (Block, et al. 2017), a lower threshold of 300 steps was set to screen out patients who may not be wearing the device based on inspection of raw data.

Given that the Althoff study uses smart phones instead of a wearable device to track steps and is likely to underestimate the true step count since the phone may not be on the user all the

time. In addition, given that patients are likely to have lower step counts therefore a cut off of 1000 steps is reasonable. Therefore if a subject has less than 1000 steps recorded in a day, we assume that device was either not worn or worn for a limited time during the day and will not be used in any calculations related to the Mhealth device. In addition, for weekly based averages we will require 4 or more valid days per week with the starting day based on the date of randomization.

9.1.7 Handling of Implausible Data

For the activity minutes of MHealth data, if implausible values are present, for example, sedentary activity minutes more than 24 hours (1440 minutes) they will be excluded from the analysis. If the sum of the sedentary, active and very active number of minutes is more than 1440 minutes then all values will be excluded from the analysis.

For the waist circumference and weight data, if a percentage change from baseline of x% in one variable is not followed by a corresponding x/4 % in the other then both would be included.

These observations will be flagged in the ADaM datasets.

9.1.8 Multiple Comparison/Multiplicity

As this study is exploratory in nature, no adjustments are made for multiplicity considerations. There will be no control of the overall type I error.

9.2 Disposition of Subjects

The disposition of subjects will be listed and summarized. The number and percent of subjects will be summarized for the following categories:

- Subjects screened
- Screen failures
- Enrolled, enrolled but not randomized
- Assigned to the Safety Set
- Randomized and further summarized for subjects:
  - Assigned to the FAS
  - Assigned to the AEECS
  - Assigned to the AEBECS
  - Completing the study
  - Discontinued (by primary reason).

For screened subjects, only the total count (overall) but no percentage will be given. For screen failures, only the total count (overall) with percentage will be given. For randomized subjects counts and percentages will be provided by study arm and overall. Subject disposition will also be summarized at Visit 14 and Visit 28 to assess dropout rate.

9.3 Demographic and Other Baseline Characteristics

All demographic and baseline characteristics data will be listed and summarized for all analysis sets.

For the FTND score a frequency table according to the following classification will be provided by analysis sets:

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

No nesting of the two grouping variables sex and cigarette consumption at V1 is planned.

9.4 Measurement of Product Compliance

There is no restriction on allocated product use.

Although subjects are being requested to use solely the product allocated to their respective study arm, it is considered that not all subjects randomized to the IQOS arms or to the cigarette arm will exclusively use the randomized product at all times during the study. Subjects may concomitantly use IQOS and cigarettes (dual use). To account for dual use of IQOS and cigarettes, adjustments for cigarette or IQOS use outside of the intended arms will be done statistically. Similarly, it is expected that subjects in the SA arm who should stay abstinent during the study may relapse to the use of tobacco or nicotine-containing tobacco products (e.g., cigarette or IQOS). In addition to the randomized arms of IQOS, Cigarette or SA , additional arms comprising of the categories Dual-1 and Dual-2 according to the definition in Section 9.1.1 ”Stratified Presentation” will be used in part of the analysis. Subjects in the ‘Others’ product use group (Figure 2) will only be listed and not analyzed.

9.5 Extent of Exposure (Product Consumption)

The number of HeatSticks, cigarettes, or/and other tobacco and nicotine-containing products used for each subject will be listed by day. Summary statistics of HeatSticks, cigarettes, or/and

other tobacco and nicotine-containing products used will be presented by month and over the entire study for the AEECS, AEBECS and FAS by study arm.

Planned Statistical Analyses

9.6.1 Endpoint Analyses

The following non-safety endpoints will be analyzed using the AEECS considering the actual study arm (i.e. product use exposure). The following sections describe the statistical models used in analyzing the endpoints as well as descriptive statistics used for summarizing the data.

In general, for model based analysis all visit data will be used and analyzed within the same model.

9.6.1.1 Statistical Analysis of Endpoints

9.6.1.1.1 To evaluate changes in VO_2max in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent

A linear mixed effects model will be used to estimate the difference in VO_2max between each arm (SA arm, IQOS-1+IQOS-2 combined arm, Dual-1+Dual-2 arm) versus cigarette arm one week after randomization (V4) and after completion (SA arm, IQOS-1 arm, IQOS-2 arm, Dual-1 arm) of the 12-week training program (V43). The degrees of freedom will be adjusted by using Kenward Roger (KR) method and the Saitterthwaite method will be used in case of convergence issues. Both V4 and V43 data will be used together in the same model. At V4, both IQOS-1 and IQOS-2 will be combined into a single IQOS group as these subjects did not attend any exercise training.

The dependent variable will be the VO_2max change from baseline with baseline VO_2max, arm, visit, sex, age and visit*arm as *fixed* covariates and subject fitted as a random effect. In addition, the following *exploratory* covariates BMI, RPC, COHb levels at each visit, cigarette consumption at V1, IPAQ score at V1, average step count 2 weeks prior to visit, self-reported daily cigarette consumption (from randomization until each visit), cigarette smoking 30-60 minutes before test for subjects in cigarette arm and IQOS use 30-60 minutes before test for subjects in IQOS arm will be assessed for adequacy in the model using a step-wise (backward) approach through the following algorithm.

1. Fit a full model using maximum likelihood with all fixed and exploratory covariates
2. Fit all possible reduced models dropping one exploratory covariate at a time.
3. Perform a likelihood ratio test for the full and reduced models.
    a. The likelihood ratio statistic for testing the adequacy of each reduced model is calculated as, $\boldsymbol{LR = \frac{lik(Reduced\ model)}{lik(Full\ model)} \sim \chi_1^2}$
4. Drop the covariate with the largest p-value for the likelihood ratio statistic if it is more than 0.05.
5. Repeat steps 2 - 4 until no more exploratory covariates can be dropped.
6. Refit the final model with the covariates remaining in the model and calculate estimates using the restricted likelihood method.

In Table 4 and Table 5 below, we show the various comparisons that will be estimated from the final fitted model. For all the groups shown under Reference Arm, estimates will be calculated within each group. Between arms comparisons are made by taking each comparator arm and comparing them against the reference arm. The corresponding 95% confidence interval for all within and between arms will also be presented. Each of these estimates will be calculated at the corresponding mean levels of a covariate. For example, the mean daily cigarette use at V43 for each arm will be used as the input instead of an overall mean daily cigarette use over all arms for comparisons in Table 5.

Table 4: Comparisons within and between arms at V4

Visit 4			
Reference Arm	Comparator Arm		
Cigarette	IQOS (IQOS-1+IQOS-2)	SA	Dual (Dual-1+Dual-2)
IQOS (IQOS-1+IQOS-2)			Dual (Dual-1+Dual-2)
SA	IQOS (IQOS-1+IQOS-2)		Dual (Dual-1+Dual-2)
Dual (Dual-1+Dual-2)			

Note: Estimates for all reference arms will be calculated, the change in $VO2_{max}$ in each comparator arm will be compared versus the reference arm.

Table 5: Comparisons within and between arms at V43

Visit 43					
Reference Arm	Comparator Arm				
Cigarette	IQOS-1		SA	Dual-1	
IQOS-1		IQOS-2	SA	Dual-1	
IQOS-2					Dual-2
SA				Dual-1	
Dual-1					
Dual-2					

An example of SAS code for the model to be used is provided below:

```
proc mixed data= _DATA_ method=ml order=data;
```

```
class SUBJECT VISIT ARM SEX;
model CHANGE = BASELINE ARM VISIT SEX AGE VISIT*ARM Covariate1 …
Covariaten/ddfm=kr;
random SUBJECT / subject = SUBJECT;
estimate "Visit 4 CC" baseline xx arm 1 0 0 0 0 0  visit 1 0 sex .5 .5 ...
CL alpha=0.05 bylevel e;
estimate "Visit 4 CC vs IQOS" baseline xx arm 1 -.5 -.5 0 0 0  visit 1 0
sex .5 .5 ... CL alpha=0.05 bylevel e;
estimate "Visit 4 IQOS" baseline xx arm 0 .5 .5 0 0 0  visit 1 0
sex .5 .5 ... CL alpha=0.05 bylevel e;
...
lsmeans VISIT*ARM / diff=control("43" "CC") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("43" "IQOS-1") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("43" " IQOS-2") CL alpha=0.05 bylevel e;
run;
```

Other covariates may be added as deemed appropriate. The same analysis will be repeated for weight-adjusted and fat-free weight adjusted VO_2max change from baseline as the dependent variable. Estimates of variability (e.g. “Covariance Parameter Estimates”) from the statistical model will be reported (at least in the original statistical output listings) and may be used to power future similar studies.

9.6.1.1.2 To evaluate changes in exercise capacity in subjects switching to IQOS (with and without training), continuing to smoke cigarettes and being smoking abstinent

The step-wise linear mixed effects model as described above for VO_2max will be used to estimate the difference in exercise capacity between the various arms versus cigarette arm one week after randomization (V4) and after completion of the 12-week training program (V43). The dependent variable will be the change from baseline in time to complete the test in seconds with baseline exercise capacity (in seconds) at V3, pre-defined work (in Calories), visit, arm, visit*arm, sex and age as fixed covariates. Subject will be fitted as a random effect. The following *exploratory* covariates BMI, RPC, COHb levels at each visit, cigarette consumption at V1, IPAQ score at V1, average step count 2 weeks prior to visit, self-reported daily cigarette consumption (from randomization until each visit), cigarette smoking 30-60 minutes before test for subjects in cigarette arm and IQOS use 30-60 minutes before test for subjects in IQOS arm will be tested as described in the previous section.

For each subject we will check that the workload for the exercise capacity test is correctly calculated based on the wattage achieved during the VO2max test at V3 times 1.25. If the workload was not calculated correct at V4 or V43, the value of the corresponding visit will be excluded from the analysis. If the value for the workload was calculated wrongly for the baseline (V3) visit but correctly calculated for subsequent visits (V4 and V43) then all values

will be excluded from this analysis. However, if the workload for all visits (V3, V4 and V43) were wrong but the same then the subject will be included in the analysis.

The comparisons of interest will follow Table 4 and Table 5 with estimates for within and between arms presented together with their 95% confidence intervals. A scatter plot of change in time to complete the test versus maximum work rate by arm will be plotted as well.

Other covariates may be added as deemed appropriate. Estimates of variability (e.g. "Covariance Parameter Estimates") from the statistical model will be reported (at least in the original statistical output listings) and may be used to power future similar studies.

A separate sensitivity analysis with all subjects including those with incorrectly calculated work load will be analyzed as well using the same analysis.

9.6.1.1.3 To assess the intensity of the exercise training in subjects participating in a training program switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

The SAS procedure PROC MIXED will be used with the restricted maximum likelihood method for estimation of the intensity of exercise training in the various (cigarette, IQOS-1 and SA) arms over time. The change from baseline (defined as V5) in exercise intensity will be the endpoint with age, visit, sex, arm, visit*arm, baseline cigarette consumption at V1 and daily cigarette consumption as covariates. Endpoints to be analyzed as exercise intensity are

- Cumulative work produced during each training session (calories and calories/kg body weight)
- Average work rate during each training session (watt and watt/kg body weight)

Estimates of the endpoints at V14, V28, V42 (corresponding to the end of the 1st, 2nd and 3rd month of training) and their 95% confidence intervals will be presented as per the comparisons listed in Table 5 except for IQOS-2 and Dual-2 groups as they did not participate in any exercise training. Estimates of study arms and pairwise comparisons of study arms will be tabulated over all visits as well as separately for each visit.

An example of SAS code for the model to be used is provided below:

```
proc mixed data= _DATA_ method=reml order=data;
class SUBJECT VISIT ARM SEX CC_CONS_GRP;
model ENDPOINT = ARM SEX AGE CC_CONS_GRP DAILY_CC_USE VISIT VISIT*ARM
/ddfm = kr;
random SUBJECT / subject = SUBJECT;
repeated VISIT / type=ar(1) subject = SUBJECT;
lsmeans ARM / diff=control("CC") CL alpha=0.05 bylevel e;
lsmeans ARM / diff=control("IQOS-1") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("VISIT#X" "CC") CL alpha=0.05 bylevel e;
```

```
lsmeans VISIT*ARM / diff=control("VISIT#X" "IQOS-1") CL alpha=0.05 bylevel
e;
run;
```

Estimates of variability (e.g. “Covariance Parameter Estimates”) from the statistical model will be reported (at least in the original statistical output listings) and may be used to power future similar studies.

The average work rate during each interval at each training session, average heart rate during each training session, time spent at various percentages of maximal work rate and time spent at various percentages of maximal heart rate will be summarized using only descriptive statistics. For the time spent at maximal work rate or heart rate the data will be summarized at V5, V14, V29 and V41 as subjects will be on training program 1 at these visits.

9.6.1.1.4 To evaluate changes in physiological parameters and perception of exertion and capacity in subjects switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

The SAS procedure PROC MIXED will be used to estimate the differences in blood composition parameters (V43) and respiratory parameters (V4 and V43) from baseline (V3) with age, sex, Ccigarette consumption at V1 and daily cigarette consumption as covariates. For the respiratory parameters, visit and visit*arm will be covariates in addition to those specified previously. Also the baseline values of the parameters (V3) will be considered as a covariate and other clinically relevant covariates of interest may be added as deemed appropriate. The dependent variable will be the change from baseline.

Endpoints to be analyzed as physiological parameters (V43) are the following blood composition parameters

- Hemoglobin mass (g)
- Red blood cell volume (mL)
- Plasma volume (mL)
- Total blood volume (mL)

Endpoints to be analyzed as respiratory parameters (V4 and V43) are the following

- Ventilation (L/min)
- Respiratory rate
- VCO2 (L/min)
- Respiratory exchange ratio (RER) (VCO2 / VO_2).

Estimates of the endpoints and their 95% confidence intervals within each arm will be reported and also differences between arms (see Table 4 for analysis related to V4 and Table 5 for V43) will be estimated and reported together with their 95% confidence intervals. Boxplots of each

physiological and respiratory endpoint listed above will be plotted versus visit and by arm within each plot.

For capillary blood lactate levels during the VO_2max test, we will fit a 4 parameter sigmoidal model for each arm at V3 and V43 respectively. Letting y_{ij} be the lactate of subject i at time j, we have

$$y_{ij} = (\beta + \beta_i) + \frac{\alpha - (\beta + \beta_i)}{1 + \exp\{\delta(\ln(W_{ij}) - \ln(\eta + \eta_i))\}} + \varepsilon_{ij}$$

where W_{ij} is the work rate for subject i at time j α and β are the minimum and maximum lactate levels respectively. The parameter η_i represents the workrate that elicits a response midway between α and β_i, and δ is a shape parameter. Lastly, the error term is modelled as ε_{ij}. Note that β_i and η_i are modelled as random effects to allow for variations in the lactate curve between subjects. An example of the sas code is as shown below, starting parameters may be found using proc nlin without the random effects. The number of random effects may be adjusted depending on model fit.

```
proc nlmixed data=_data_;
   parms b0=xx e0=xx a=xx d=xx s2=xx s2b1=xx s2b2=xx cb12=xx;
   beta   = b0 + bi;
   eta    = e0 + ei;
   pred = beta+(a-(beta))/(1+exp(d(ln(w)-ln(eta))));
   model vo2    ~ normal(pred,s2);
   random bi ei ~ normal([0,0],[s2b1,cb12,s2b2]) subject=subject;
run;
```

Using the fitted models, we will back-calculate the value of W*(conditional on β_i and η_i) that elicits a response of y=4 mmol (Heck 1985) using the following formula for each arm and visit.

$$W^* = \hat{\eta}\left(\frac{\hat{\alpha} - 4}{4 - \hat{\beta}}\right)^{\frac{1}{\hat{\delta}}}$$

Note that $\hat{\alpha}$, $\hat{\beta}$, $\hat{\delta}$ and $\hat{\eta}$ are the respective estimates from the fitted model. Within each arm, we will present the change in workload at a lactate level of 4mmol from V3 to V43. The mean curve estimated from the model for Cigarette, IQOS-1, IQOS-2 and SA will be plotted versus workload for each arm and visit, in these plots a horizontal line at 4 mmol will be drawn as well. The same plot will be done for Cigarette, IQOS-1, IQOS-2, Dual-1 and Dual-2.

For the endpoints VO_2 and heart rate (V4 and V43), a linear mixed effects model will be used with the restricted maximum likelihood method for estimation of the parameter in the various (cigarette, IQOS-1 and SA) arms versus workload. Each endpoint will be fitted with workload, age, visit, sex, arm, visit*arm, baseline cigarette consumption at V1 and daily cigarette

consumption as covariates. Additional covariates may be added as appropriate based on the results of the step wise modelling in section 9.6.1.1.1.

The mean curve estimated from the model for the CIGARETTE, IQOS-1, IQOS-2 and SA arm will be plotted versus workload for each arm and visit. The curves will be plotted for CIGARETTE, IQOS-1, IQOS-2, Dual-1 and Dual-2 in one plot as well.

An example of SAS code for the model to be used is provided below:

```
proc mixed data= _DATA_ method=reml order=data;
class SUBJECT VISIT ARM SEX CC_CONS_GRP;
model ENDPOINT = WORKLOAD ARM SEX AGE CC_CONS_GRP DAILY_CC_USE VISIT
VISIT*ARM /ddfm = kr;
random SUBJECT / subject = SUBJECT;
run;
```

Estimates of variability (e.g. “Covariance Parameter Estimates”) from the statistical model will be reported (at least in the original statistical output listings) and may be used to power future similar studies.

9.6.1.1.5 To monitor trends of daily physical activity levels during the study in subjects switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

The SAS procedure PROC MIXED will be used with the restricted maximum likelihood method for evaluation of the aggregated data for step counts from the mobile health (mHealth) wearable in the study arms over time. For step count, the dependent variable will be the average step count by month, this is derived by averaging the daily step count from V4 to V14 (Month 1), V14 to V28 (Month 2) and V28 to V43 (Month 3) divided by the number of valid days during each defined period. For a definition of valid day for step count refer to section 9.1.6.

The variables Arm, Visit, Visit*Arm, Sex, age, and baseline cigarette consumption at V1, baseline step count (see section 9.1.3) will be considered as covariates, subject will be used as a random effect. The Kenward-Roger adjustment for degrees of freedom will be use.

Estimates of the monthly step count and their 95% confidence intervals within each arm will be reported and also between arms as per Table 5.

An example of SAS code for the model to be used is provided below:

```
proc mixed data= _DATA_ method=reml order=data;
class SUBJECT MONTH ARM SEX CC_CONS_GRP;
model ENDPOINT = ARM SEX AGE CC_CONS_GRP MONTH MONTH*ARM /ddfm = kr;
random SUBJECT / subject = SUBJECT;
lsmeans ARM / diff=control("CC") CL alpha=0.05;
lsmeans ARM / diff=control("IQOS-1") CL alpha=0.05;
```

```
lsmeans ARM / diff=control("IQOS-2") CL alpha=0.05;
lsmeans MONTH*ARM / diff=control("MONTH#X" "CC") CL alpha=0.05;
lsmeans MONTH*ARM / diff=control("MONTH#X" "IQOS-1") CL alpha=0.05;
lsmeans MONTH*ARM / diff=control("MONTH#X" "IQOS-2") CL alpha=0.05;
run;
```

The percentage of time active and percentage of time very active will be summarized on a monthly basis descriptively.

For each subject, average step counts will be plotted at baseline (V2-V3), Months 1 to 3. These will be presented as boxplots with step count on the y-axis and time on the x-axis for each product use group.

In addition, moving averages using a 7-day time window for daily step counts for all subjects will be plotted as well. If data is missing on one or more days, the average step count will be divided by the number of valid days available over the 7 day period. These will be plotted for all subjects versus day and color coded by product use arm. A separate plot will be created by averaging the daily moving average for each subject within a product use arm and plotted against time as well.

9.6.1.1.6 To describe changes in biological health markers in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

The SAS procedure PROC MIXED will be used with the restricted maximum likelihood method for evaluation of the biological health markers at V14, V28 and V43. The dependent variable will be the change from baseline of the biological health markers with age, sex, arm, visit, visit*arm, baseline cigarette consumption at V1 and daily cigarette consumption as covariates. Subject will be fitted as a random effect.

The baseline values of the parameters (V3) will be considered as covariate and other clinically relevant covariates of interest may be added as deemed appropriate. Estimates of the endpoints and their 95% confidence intervals within each arm will be reported and also between arms as per Table 5.

An example of SAS code for the model to be used is provided below:

```
proc mixed data= _DATA_ method=reml order=data;
class SUBJECT VISIT ARM SEX CC_CONS_GRP;
model CHANGE = BASELINE ARM SEX AGE CC_CONS_GRP DAILY_CC_USE VISIT
VISIT*ARM /ddfm = kr;
random SUBJECT / subject = SUBJECT;
lsmeans VISIT*ARM / diff=control("14" "CC") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("14" "IQOS-1") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("14" "IQOS-2") CL alpha=0.05 bylevel e;
```

```
lsmeans VISIT*ARM / diff=control("28" "CC") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("28" "IQOS-1") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("28" "IQOS-2") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("43" "CC") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("43" "IQOS-1") CL alpha=0.05 bylevel e;
lsmeans VISIT*ARM / diff=control("43" "IQOS-2") CL alpha=0.05 bylevel e;
run;
```

9.6.1.1.7 To describe changes in weight, body fat and waist circumference in subjects switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

The SAS procedure PROC MIXED as described above for biological health markers will be used to evaluate the change from baseline in body fat and waist circumference (V4 and V43) with age, sex, visit, product use, visit*product use and daily cigarette consumption as covariates. Subject will be fitted as a random effect. Also the baseline values of the parameters (V3) will be considered as covariate and other clinically relevant covariates of interest may be added as deemed appropriate. Estimates within and between groups similar to those shown in Table 4 and Table 5 will be provided.

For body weight, the model presented in section 9.6.1.1.3 will be used to assess the changes at V14, V28 and V43 with change from baseline (V3) as the endpoint and age, sex, visit, product use, visit*product use and daily cigarette consumption as covariates. Subject will be fitted as a random effect. Also the baseline values of the parameters (V3) will be considered as covariate and other clinically relevant covariates of interest may be added as deemed appropriate. Estimates within and between groups at V14, V28 and V43 will be calculated together with their respective 95% confidence intervals.

Box plots for the change from baseline in weight for the study arms by month will be presented. For body fat and waist circumference we will present the study arms side by side separately for each visit (V4 and V43).

9.6.1.1.8 To monitor levels of exposure to CO, nicotine, nitrosamines and acrylonitrile in subjects switching to IQOS, continuing to smoke cigarettes and being smoking abstinent

The corresponding endpoints, i.e. the exhaled CO and COHb% in blood, NEQ, NNAL and CEMA in urine (the latter three adjusted for creatinine) will be summarized descriptively, see Section 9.6.1.2 .

9.6.1.2 Descriptive Analysis and Listing of Endpoints

Non-safety endpoints will be listed and summarized for the AEECS using the study arm and by visit or by month in the case of more densely collected endpoints. See Section 9.1.2 “Descriptive Statistics” for a summary of continuous and categorical variables.

In general, the descriptive statistics for the endpoints will be grouped by the following arms, Group1: CIGARETTE, IQOS-1, IQOS-2 and SA. Another set of descriptive statistics Group 2: Dual-1, Dual-2 and Others will be also presented to manage the size of the tables.

The endpoints related to the VO2max (including RPC; respiratory parameters; heart rate and oxygen uptake at VO2max) and exercise capacity test will have baselines dependant on the product use group for the Week 1 (acute) and Overall (training) analysis period. These baselines will be distinguished as Acute Baseline and Training Baseline respectively. All other endpoints will have their baseline dependant on the product use Overall.

9.6.1.2.1 VO_2max

For VO_2max recorded values and the changes from baseline will be listed and summarized. For each visit within each study arm VO_2max (observed values, weight-adjusted- and fat free weight-adjusted values) will be summarized by sex and cigarette consumption at V1. Both acute and training baseline will be presented for $VO2_{max}$.

Box plots will be prepared for VO_2max changes from baseline presenting the study arms side by side separately for each visit.

9.6.1.2.2 Exercise capacity

For exercise capacity recorded values and the changes from baseline will be listed and summarized. For each visit within each study arm exercise capacity will in addition be summarized by sex and cigarette consumption at V1. Both acute and training baseline will be presented for exercise capacity.

Box plots will be prepared for changes in exercise capacity from baseline presenting the study arms side by side separately for each visit.

9.6.1.2.3 Intensity of exercise training

For the intensity of the exercise training recorded values will be listed. The following variables will be summarized at V5 (baseline), V14, V29 and V41 so that all subjects are on training program 1.

- Cumulative work produced during each training session (calories and calories/kg body weight)
- Average work rate during each training session (watt and watt/kg body weight)
- Average heart rate (HR) during each training session (bpm)

For cumulative work and average work rate they will be plotted for V5, V14, V29 and V41 using box plots with the response on the y-axis and month on the x-axis, within each month there will be a box plot for each product use category. Average heart rate will only be summarized and not plotted as the exercises are controlled for heart rate.

- Time spent at 0-50%, 50-65%, 65-75%, 75-90%, >90% of maximal work rate during each training session (sec)
- Time spent at 0-50%, 50-65%, 65-75%, 75-90%, >90% of maximal HR during each training session (sec).

For the time spent at maximum work rate or heart rate the data will again be summarized at only visits V5, V14, V29 and V41. Time spent at the various intervals of maximal work rate will be plotted using stacked bars for each product use group at each visit. Time spent in an interval of maximal heart rate will not be plotted.

Within each study arm the following variables will be summarized by interval (0-5, 6-20, 21-30, and 31-40 minutes) at V5, V14, V29 and V41.

- Intensity (%) {percent of maximal workload}
- Work rate (watt)
- Work rate per body weight (watt/kg body weight).

9.6.1.2.4 Physiological parameters and perception of exertion and capacity

For blood composition recorded values and derived (or computed) values as well as the changes from baseline will be listed and commonly summarized by arm. Derived endpoints (Hbmass, RBCV, BV, and PV) are described in Section 5.2 “Blood composition as determined by CO re-breathing method”. Lactate values (continuous) and Borg RPE scale values (categorical) will be listed; they will be summarized within each study arm (and overall) for each visit at rest and at each 25 Watt workload increase.

For the BORG scale, a mosaic plot showing the number of subjects with BORG scale > 16 and the number of subjects with BORG scale < 16 (y-axis) versus maximum wattage during the VO2max test (x-axis) will be plotted separately for each study arm.

Respiratory parameters (continuous) with changes from baseline and Rating of Perceived Capacity (RPC) scale values (categorical) will be listed and summarized. Both acute and

training baseline will be presented for these parameters as well. Although the RPC is a categorical (actually ordinal) scale it will be summarized as a continuous variable by visit, arm and stratified by sex. Scatter plots of the RPC scale vs the corresponding VO_2max value at each visit will be plotted.

Heart rate and oxygen uptake (VO_2) will be listed; these data will be summarized within each study arm for each visit at rest and at each 25 Watt workload increase.

9.6.1.2.5 Daily physical activity

For step count raw values will be averaged and reported by month. Measures of physical activity and changes from baseline will be listed and summarized by visit and actual product use for baseline and each month.

9.6.1.2.6 Biological health markers

For biological health markers the changes from baseline for each arm as endpoint, therefore the observed values and the changes from baseline will be listed and summarized by visit and arm.

9.6.1.2.7 Weight, body fat and waist circumference

For body fat and waist circumference (V3, V4 and V43) both the observed values and the changes from baseline will be listed and summarized by visit and arm. For weight (V3-V43) the data will be summarized for each month (see section 9.1.3.1) by visit and arm.

9.6.1.2.8 Exposure to CO, nicotine, nitrosamines and acrylonitrile

For exhaled CO and COHb% in blood both the observed values and the changes from baseline will be listed and summarized at Baseline, V4 and by month (see section 9.1.3.1) for each arm. The reporting of COHb% will on the log-scale then back-transformed to obtain the geometric means.

For NEQ, NNAL and CEMA in urine the values adjusted for creatinine will be listed. For the latter also the changes from baseline will be listed. The values adjusted for creatinine and their changes from baseline will be summarized by visit and arm. Similar to COHb%, these values will be log-transformed and means calculated, these values are then back transformed to obtain the geometric means.

9.6.1.2.9 Self-reported nicotine and tobacco containing product use

For self-reported daily cigarette consumption and/or IQOS use (self-reported nicotine and tobacco containing product use) see Section 9.5 “Extent of Exposure (Product Consumption)”.

9.6.1.3 Supportive/Sensitivity Analyses

If deemed appropriate, selected endpoint analyses will additionally be performed for the FAS.

9.6.2 Exploratory Analyses

9.6.2.1 Effect of subject preference on compliance

To explore if a subject’s preference has any influence on the compliance of subjects to product use, a log-linear model will be used to estimate the dependence of subject compliance based on randomized arm and preference.

For these subjects, the number of subjects who were compliant (actual arm = randomized arm) and not compliant will be tabulated and used as the response (1=Compliant, 0=Not Compliant). The SAS procedure CATMOD will be used to analyze this data and the sample code is as provided below.

```
proc catmod data=_data;
   response 1 0;
   weight Count;
   model Compliance=Preference|Randomized_Arm / freq prob;
   title2 'Saturated Model';
run;
```

A contingency table showing the compliance vs whether a subject was randomized to preferred group will be presented as well.

9.6.2.2 To evaluate changes in VO_2max in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent based on AEBECS

The impact of biochemical verification of subjects’ smoking status (Section 8.4) for VO_2max will be assessed using the same approach as the analysis described in Section 9.6.1.1.1 except that we will use the AEBECS analysis set instead of AEECS.

Descriptive statistics for this analysis set will be presented as well using the same method as described in Section 9.6.1.2.1.

9.6.2.3 To evaluate changes in exercise capacity in subjects switching to IQOS (with and without training), continuing to smoke cigarettes and being smoking abstinent - AEBECS

The impact of biochemical verification of subjects' smoking status (Section 8.4) for exercise capacity will be assessed using the same approach as the analysis described in Section 9.6.1.1.2 except that we will use the AEBECS analysis set instead of AEECS.

Descriptive statistics for this analysis set will be presented as well using the same method as described Section 9.6.1.2.2.

9.6.2.4 To describe changes in biological health markers in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

The impact of biochemical verification of subjects' smoking status (Section 8.4) for biological health markers will be assessed using the same approach as the analysis described in Section 9.6.1.1.6 except that we will use the AEBECS analysis set instead of AEECS.

Descriptive statistics for this analysis set will be presented as well using the same method as described Section 9.6.1.2.6.

9.6.2.5 To describe changes in levels of exposure to CO, nitrosamines and acrylonitrile in subjects switching to IQOS (with and without training), continuing to smoke cigarettes (with training) and being smoking abstinent (with training)

The impact of biochemical verification of subjects' smoking status (Section 8.4) for levels of exposure to CO, nitrosamines and acrylonitrile will be assessed using descriptive statistics for this analysis set as described Section 9.6.1.2.8.

9.6.3 Additional Analyses

Not Applicable.

9.6.4 Safety Evaluation

9.6.4.1 Adverse Events

AEs (including SAEs) will be collected from the time of ICF signature until the EOS for all enrolled subjects. Information recorded will include: verbatim description of the AE/SAE, start and stop dates, seriousness, severity (intensity), relatedness (to IP and study procedures), expectedness, action taken (e.g., whether or not the AE/SAE led to the subject's withdrawal from the study), and outcome (e.g., resolved, stabilized).

Adverse events which occur during the screening period will be listed together with corresponding information.

Adverse events (AEs) will be coded using the MedDRA® dictionary.

A summary showing the overall number of AEs, SAEs, severity, expectedness, relatedness, AE leading to death, AE leading to action taken and treatment given for an AE will be presented.

The frequency of AEs and number (and percentage compared to total) of subjects for each AE will be determined within each study arm and for all study arms. This will be done for a given AE with respect to the system organ class (SOC) and preferred term according to MedDRA® will be determined within each study arm and for all study arms (total) and tabulated. Tabulation will be sorted by the decreasing frequencies within the system organ classes and preferred term (within the system organ class).

The frequencies of the severity (intensity) and relationship to study procedures and IP, expectedness, for each adverse event will be tabulated for each study arm by SOC and preferred term.

9.6.4.1.1 Serious Adverse Events (Including Deaths)

The frequency of SAEs will be determined within each study arm and for all study arms (total). Furthermore, the absolute and relative frequencies for subjects with a given SAE with respect to the system organ class (SOC) and preferred term according to MedDRA® will be determined within each study arm and for all study arms (total) and tabulated. Tabulation will be sorted by the decreasing frequencies within the system organ classes and preferred term (within the system organ class).

9.6.4.1.2 Adverse Events Leading to Discontinuation

The frequency of AEs/SAEs leading to the subject's withdrawal from the study will be determined within each study arm and for all study arms (total) and summarized.

9.6.4.2 Device Events

All IQOS device issues and/or malfunctions will be listed, including event description, severity of event, AE relationship, proposed solution and start/stop dates/times.

The frequency of device events will be determined within each study arm and for all study arms (total). Furthermore, the absolute and relative frequencies for subjects with a given device event with respect to the system device event term will be determined within each study arm and for all study arms (total) and tabulated.

The frequencies of the severity (intensity) and relationship to study procedures for each device events will be tabulated for each study arm by device event term.

A summary table of AEs related and expected to device events will be presented as well.

9.6.4.3 Clinical Laboratory Evaluation

The grading scheme used in the Common Terminology Criteria for Adverse Events and Common Toxicity Criteria ([CTCAE] version 4.03) will be used. These CTCAE grades (1, 2, 3, 4, if applicable) will be documented in the CRF.

Individual values of the clinical laboratory parameters (clinical chemistry, hematology, and urinalysis for safety panel) will be listed including reference range, flags for values outside the reference range (H above the reference range or L below the reference range), CTCAE grades (only 1, 2, 3, or 4), interpretation and change from baseline. Additionally, all values outside the reference range will be listed by parameter.

Descriptive statistics will be provided for continuous laboratory data (observed values and changes from baseline) by study arm and for all study arms (total) and if appropriate, by scheduled time (e.g. visit).

A frequency table will show the number and percentage of subjects below lower limit, within the reference range and above the upper limit for each parameter. This table will be done by study arm and for all study arms (total) and if appropriate, by scheduled time (e.g. visit). CTCAE grades may also be summarized using a frequency table.

The individual data of serology, drug screen, alcohol breath test, and pregnancy tests will be listed only.

9.6.4.4 Medications, Physical Findings, Vital Signs and Other Observations Related to Safety

9.6.4.4.1 Medical History and Concomitant Disease

Medical history is defined as any condition that started and ended prior to ICF signature. A concomitant disease is defined as any condition that started prior to ICF signature and is still ongoing at V1. Medical history and concomitant disease will be coded using the MedDRA® dictionary. The data will be summarized by SOC and PT for each study arm and also listed.

9.6.4.4.2 Prior and Concomitant Medication

All medication taken within 4 weeks prior to the ICF signature will be considered prior medication. All medication taken from screening visit to end of the safety follow-up period will be considered concomitant medication. Medication will be coded using the WHO Drug Global dictionary. The data will summarized at ATC Levels 1 and 2 by preferred drug name, study arm and presented in a listing.

9.6.4.4.3 Physical Examination

A full physical examination will include review of general appearance, skin, head, eyes, ears, nose and throat, thyroid gland, chest, lungs, back, abdomen, dentition, cardiovascular, gastrointestinal, musculoskeletal and neurological systems.

Results (normal, abnormal not clinically significant, abnormal clinically significant) and abnormalities - if 'Abnormal clinically significant' – will be listed.

A frequency table will show the number and percentage of subjects of the different results. This table will be done by study arm and for all study arms (total) and by scheduled time (e.g. visit).

9.6.4.4.4 Vital Signs

Systolic and diastolic blood pressure, pulse rate and respiratory rate will be recorded. Observed values and the changes from baseline will be listed and summarized by study arm and for all study arms (total) and by scheduled time (e.g. visit).

9.6.4.4.5 Spirometry

FEV1, FEV1 % predicted, FVC, FVC % predicted, and FEV1/FVC will be recorded. The screening visit (V1) is used as baseline. Observed values and the changes from baseline will

be listed and summarized by study arm and for all study arms (total) and by scheduled time (e.g. visit).

9.6.4.4.6 Electrocardiogram

Heart rate (HR), PR, QRS, QT, QTcB, and QTcF intervals will be recorded. Observed values and the changes from baseline will be listed and summarized by study arm and for all study arms (total) and by scheduled time (e.g. visit).

9.6.4.4.7 Assessment of Cough

VAS values and changes from baseline as well as scores of Likert scales measuring the intensity and frequency of cough, and the amount of sputum production will be listed for all subjects who filled in the questionnaires. VAS will be summarized by means of descriptive statistics; Likert scales will be summarized by means of frequency tables by study arm and for all study arms (total) and by scheduled time (e.g. visit).

The number and percentage of subjects reporting no cough and reporting a cough will be provided.

10 ANALYSES AND REPORTING

10.1 Interim Analyses and Data Monitoring

No interim analysis is planned for this study.

The Clinical Research Associate (“Monitor”) from CRS will be responsible for the monitoring of the study. Monitoring will be performed according to CRS’s Standard Operating Procedures (SOPs) and as per the agreed monitoring plan with the Sponsor.

10.2 Safety Reporting

Not applicable.

10.3 Topline Results

The topline TFLs are flagged in Section 13.2 “Tables, Figures & Listings”.

10.4 Final Analyses

The final analysis will be performed after database lock and finalization of this SAP. A data review meeting will be held prior to database lock.

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 unplanned analyses conducted prior to the finalization of the CSR will be clearly documented in the CSR. Any additional analysis performed after the finalization of the CSR will either be included in an amendment of the CSR or be created as an addendum to the CSR.

10.5 ClinicalTrials.Gov Reporting

The TFLs which will be published on the Clinical trials.gov website are flagged in Section 13.2 “Tables, Figures & Listings”.

11 DATA PRESENTATION

<Enter Text>

12 REFERENCES

Albrecht, AE, BH Marcus, M Roberts, DE Forman, and AF Parisi. 1998. "Effect of smoking cessation on exercise performance in female smokers participating in exercise training." *American Journal of Cardiology* 82 (8): 950-5.

Althoff, Tim, Rok Sosič, Jennifer L. Hicks, Abby C. King, and Scott L. Leskovec, Jure Delp. 2017. "Large-scale physical activity data reveal worldwide activity inequality: Supplementary Information." *Nature* 547: 336–339.

Berg, CJ, GL Schauer, JS Ahluwalia, and NL Benowitz. 2012. "Correlates of NNAL levels among nondaily and daily smokers in the college student population." *Current Biomarker Findings* 2: 87-94.

Block, VJ, Lizée A, Crabtree-Hartman E, Bevan CJ, Graves JS, Bove R, Green AJ, et al. 2017. "Continuous daily assessment of multiple sclerosis disability using remote step count monitoring." *Journal of Neurology* 264 (2): 316-326.

Borg, GA. 1982. "Psychophysical bases of perceived exertion." *Medicine & Science in Sports & Exercise* 14 (5): 377-81.

Claussen, A, A Donelli, W Copalu, N Lama, R Weitkunat, C Haziza, and F Lüdicke. 2019. "Exploring Discriminant Capacity of Urinary CEMA as Combustion Marker in Tobacco Users - A Population Pharmacokinetic Approach." *SRNT.*

Heatherton, TF, Kozlowski LT, Frecker RC, and Fagerström KO. 1991. "The Fagerström test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire." *British Journal of Addiction* 86 (9): 1119-27.

Heck, H., Mader, A., Hess, G., Mücke, S., Müller, R., & Hollmann, W. 1985. "Justification of the 4-mmol/l lactate threshold." *International Journal of Sports Medicine* 06 (03): 117-130.

IPAQ. 2005. *IPAQ scoring protocol.* Nov. Accessed FEB 2020. https://sites.google.com/site/theipaq/scoring-protocol.

Siebenmann, C., S. Keiser, P. Robach, and C. Lundby. 2017. "CORP: the assessment of total hemoglobin mass by carbon monoxide rebreathing." *Journal of Applied Physiology* 123 (3): 645-654.

Wisen, AG, RG Farazdaghi, and B Wohlfart. 2002. "A novel rating scale to predict maximal exercise capacity." *European Journal of Applied Physiology* 87 (4-5): 350-7.

13 APPENDICES

13.1 Study Assessments

Table 13–1: SCHEDULE OF EVENTS

	Screening	Familiarization	Baseline	Exercise Tests Acute Effect	Training Visits	Medical Visit	Training Visits	Medical Visit	Training Visits	Exercise Tests Training Effect	Discharge / Early Termination	28-day Safety follow-up
Visit	**1**	**2**	**3**	**4**	**5-14**	**14**	**15-28**	**28**	**29-42**	**43**	**44**	
Informed consent[3]	•											
Inclusion/Exclusion	•	•[4]										
Advice on the risks of smoking/smoking cessation advice and debriefing on IQOS	•	•				•		•			•• - if ET	
Demographics	•											
Fagerstrom's test for nicotine dependence (FTND)			•									
Tobacco and nicotine containing product use history and habits	•											
Cotinine urine test	•	•										
Exhaled CO	•	•	•	•	•	•	•	•	•	•		
International Physical Activity Questionnaire (IPAQ-short)	•											
Urine drug screen[5]	•	•										
Alcohol breath test	•	•										
HIV, Hepatitis B and C	•											
Safety laboratory hematology, clinical chemistry and urinalysis	•		•			•		•			•• - if ET	

[3] ICF for study participation including section for optional participation in sub-study interviews, and separate ICF for optional biobanking.

[4] Urine pregnancy test, cotinine urine test, exhaled CO, alcohol breath test, urine drug screen will be performed at V2 before enrollment.

[5] Amphetamines, barbiturates, benzodiazepines, cannabinoids, cocaine, opiates, methadone, methamphetamines, phencyclidine, tricyclic antidepressants,.

	Screening	Familiarization	Base-line	Exercise Tests Acute Effect	Training Visits	Medical Visit	Training Visits	Medical Visit	Training Visits	Exercise Tests Training Effect	Discharge / Early Termination	28-day Safety follow-up
Visit	**1**	**2**	**3**	**4**	**5-14**	**14**	**15-28**	**28**	**29-42**	**43**	**44**	
Cough assessment questionnaire (VAS scale, and 3 Likert scales)			•			•		•			•	
Urine pregnancy test	•	•	•		•[6]	•		•			•• - if ET	
Physical examination	•		•			•		•			•• - if ET	
ECG	•		•			•		•			•• - if ET	
Spirometry	•										•• - if ET	
Vital signs[7]	•	•	•	•		•		•		•	•• - if ET	
Height	•											
Weight	•		•	•	•	•	•	•	•	•		
BMI	•											
Medical history and concomitant diseases	•											
Concomitant disease status		•	•	•	•	•	•	•	•	•	•	
AE/SAE recording	•	•	•	•	•	•	•	•	•	•	•• - if ET	•[8]
Prior and concomitant medications	•	•	•	•	•	•	•	•	•	•	•• - if ET	•[9]
Lifestyle questionnaire			•									
Rating of Perceived Capacity (RPC) scale			•	•						•		

[6] Urine pregnancy test will not be performed at each training but at V5 only.

[7] Systolic and diastolic blood pressure, respiratory rate, pulse rate.

[8] Spontaneous reporting of new AEs and active follow-up of AEs ongoing at V44.

[9] Only concomitant medications used for treatment of AE/SAEs ongoing at discharge at V43.

	Screening	Familiarization	Base-line	Exercise Tests Acute Effect	Training Visits	Medical Visit	Training Visits	Medical Visit	Training Visits	Exercise Tests Training Effect	Discharge / Early Termination	28-day Safety follow-up
Visit	**1**	**2**	**3**	**4**	**5-14**	**14**	**15-28**	**28**	**29-42**	**43**	**44**	
Product preference[10]			•									
Smoke a cigarette/Heatstick before VO_2max test/exercise training[11]			•	•	•		•		•	•		
Snack and water after VO_2max test	•		•	•						•		
Bike ergometer – VO_2max test[12]	•		•	•						•		
Readiness to comply with study procedures	•											
Enrollment		•										
Bike ergometer – Exercise capacity[13]		•	•	•						•		
mHealth wearable demonstration	•											
mHealth wearable distribution and training		•										
mHealth wearable return										•	• - if ET	
Randomization			•									
Demonstration of IQOS	•											

[10] Will be asked before randomization. As an instruction to answer this question, subjects will be informed that their response will not influence the randomization process.

[11] As per randomized arm before exercise tests and trainings. Only applies for exercise tests for subjects in *IQOS-2* arm.

[12] The VO_2max test will be performed on a bike ergometer using a mouthpiece and nose clip. Measurements will begin with the subjects resting for 3 min seated on the bike ergometer whereafter the workload will be increased by 25W every 60 sec until volitional fatigue. Subjects should keep cadence between 60 and 90 rpm during the test. At V3, V4 and V43, heart rate and VO_2 will be recorded at rest and at each 25W increase.

[13] The time to complete a pre-defined work as fast as possible on a bike ergometer. The criteria for completion of the test is the completion of the pre-defined work determined as 25% more work than was completed during the VO_2max test at V1 for the test at V2, and at V3 for the tests at V3, V4 and V44.

	Screening	Familiarization	Base-line	Exercise Tests Acute Effect	Training Visits	Medical Visit	Training Visits	Medical Visit	Training Visits	Exercise Tests Training Effect	Discharge / Early Termination	28-day Safety follow-up
Visit	**1**	**2**	**3**	**4**	**5-14**	**14**	**15-28**	**28**	**29-42**	**43**	**44**	
IQOS starter kit distribution and device return			•							•	•- if ET	
Product use diary Distribution/Return		•								•	• - if ET	
Bike ergometer 40-min training session[14]					•		•		•			
Body fat percentage			•	•						•		
Waist circumference			•	•						•		
Blood composition and volume (CO re-breathing method) including blood sample for hematocrit Hb and COHb before CO-rebreathing, and COHb after CO-rebreathing.			•							•		
Perceived exertion - RPE Borg scale [15]			•							•		
Blood COHb%			•	•		•		•		•		
Capillary blood lactate levels[16]			•							•		

[14] Training sessions will be performed according to Appendix 2 and will be adjusted according to the subject's resting and maximal heart rate.

[15] Rate of perceived exertion will be measured at rest and at every 25 W increase during the VO_2 max test until the end of the test.

[16] Capillary (fingertip or earlobe) blood samples will be obtained at rest and at each workload and analysed for lactate. Blood collection should be started 40-45 sec into the workload and no blood sample should be obtained at maximal effort. The final blood sample should be obtained 60 sec after termination of the bike exercise.

	Screening	Familiarization	Base-line	Exercise Tests Acute Effect	Training Visits	Medical Visit	Training Visits	Medical Visit	Training Visits	Exercise Tests Training Effect	Discharge / Early Termination	28-day Safety follow-up
Visit	**1**	**2**	**3**	**4**	**5-14**	**14**	**15-28**	**28**	**29-42**	**43**	**44**	
Blood sampling for biological health markers[17]			•			•		•		•		
Spot urine collection for urinary exposure markers (NEQ, NNAL, CEMA, creatinine)[18]			•	•		•		•		•		
Biobanking sample collection– blood, urine, saliva			•			•		•		•		
Biobanking sample collection- sweat			•							•		
Qualitative interview on health and functioning[19]			•							•		
Discharge											•	

[17] Blood sampling will be performed on fasting state (at least 8 hours of fasting).

[18] Urine collection to be performed as late as possible during the day.

[19] Subjects that have signed the optional section of the ICF for the sub-study on health on functioning will have a one-hour, qualitative interview over the phone or computer between V2 and V3, and in the end of the study between V42 and V43.

13.2 Tables, Figures & Listings

Section or Title	Topline	Dry Run
15.1. Figures		
Figure 15.1.1.1: Box Plot of VO2max change from baseline – AEECS	X	X
Figure 15.1.2.1: Box Plot of Exercise Capacity Change from Baseline – AEECS	X	X
Figure 15.1.3.1.: Boxplot of Average Work Rate during Training Sessions by Month and Product Use Group – AEECS		X
Figure 15.1.3.2.: Boxplot of Cumulative Work Produced during Training Sessions by Month and Product Use Group – AEECS		
Figure 15.1.3.3.: Boxplot of Average Work Rate in Each Interval during Training Session by Month and Product Use Group – AEECS		
Figure 15.1.3.4.: Time Spent in an interval of Maximal Work Rate during each Training Session at V5, V14, V29 and V41 – AEECS		X
Figure 15.1.4.1.: Fitted Blood Lactate versus Work Rate during the VO2max Test by Visit for CIGARETTE, IQOS-1, IQOS-2 and SA – AEECS		X
Figure 15.1.4.2.: Fitted Blood Lactate versus Work Rate during the VO2max Test by Visit for CIGARETTE, IQOS-1, IQOS-2, Dual-1 and Dual-2 – AEECS		
Figure 15.1.4.3.: Heart Rate Curve versus Work Rate during the VO2max Test for CIGARETTE, IQOS-1, IQOS-2 and SA – AEECS		
Figure 15.1.4.4.: Heart Rate Curve versus Work Rate during the VO2max Test for CIGARETTE, IQOS-1, IQOS-2, Dual-1 and Dual-2 – AEECS		
Figure 15.1.4.5.: Oxygen Uptake Curve versus Work Rate during the VO2max Test for CIGARETTE, IQOS-1, IQOS-2 and SA – AEECS		
Figure 15.1.4.6.: Oxygen Uptake Curve versus Work Rate during the VO2max Test for CIGARETTE, IQOS-1, IQOS-2, Dual-1 and Dual-2 – AEECS		
Figure 15.1.4.7.: Physiological (Blood Composition) Parameters vs Visit by Arm – AEECS		X
Figure 15.1.4.8: Respiratory Parameters vs Visit by Arm – AEECS		
Figure 15.1.4.9.: Mosaic Plot of Borg Rating (>16 and ≤16) of Perceived Exertion Scale versus maximum workload during the VO2max Test by Visit – AEECS		X
Figure 15.1.4.10.: Scatter Plot of RPC versus VO2max by Visit – AEECS		
Figure 15.1.5.1.: Boxplot of Average Step Count versus month by Arm - AEECS		
Figure 15.1.5.2.: Moving Average (7-Day Window) Plot of Daily Step Count versus month by Subject - AEECS		
Figure 15.1.5.3.: Moving Average (7-Day Window) Plot of Daily Step Count versus month by Arm - AEECS		

Figure 15.1.6.1: Box Plot of Weight Change from Baseline by Month – AEECS		
Figure 15.1.6.2: Box Plot of Body Fat Change from Baseline – AEECS		
Figure 15.1.6.3: Box Plot of Waist Circumference Change from Baseline – AEECS		
15.2 Tables		
Table 15.2.1.1: Summary of Subject Disposition - Screened Subjects		X
Table 15.2.1.2: Summary of Reasons for Discontinuation – Subjects in FAS, AEECS and AEBECS		
Table 15.2.1.3: Analysis Sets and Summary of Reason for Exclusion from Analysis Sets		
Table 15.2.1.4: Number of Subjects in the Analysis Sets by Arm – AEECS, AEBECS, Full Analysis Set		
Table 15.2.1.3.1: Demographic and Baseline Characteristics – AEECS	X	X
Table 15.2.1.3.2: Demographic and Baseline Characteristics – AEBECS		
Table 15.2.1.3.3: Demographic and Baseline Characteristics – Full Analysis Set		
Table 15.2.1.4.1: Summary of Product Use (Week 1) – AEECS	X	X
Table 15.2.1.4.2: Summary of Product Use (Overall) – AEECS		
Table 15.2.1.4.3: Summary of Product Use (Week 1) – AEBECS		
Table 15.2.1.4.4: Summary of Product Use (Overall) – AEBECS		
Table 15.2.1.4.5: Summary of Product Use (Overall) – Full Analysis Set		
Table 15.2.1.5: Summary of Product Preference – AEECS		
Table 15.2.1.6.1: Summary of Protocol Deviations – Safety Set		
Table 15.2.1.7.1: Summary of Lifestyle Questionnaire - AEECS		
Table 15.2.1.8: Summary of IPAQ Questionnaire – AEECS		
Table 15.2.2.1.1: Analysis of VO2max Change from Baseline (Acute Effect, V4) – AEECS	X	X
Table 15.2.2.1.2: Analysis of VO2max Change from Baseline (Training Effect, V43) – AEECS	X	X
Table 15.2.2.1.3: Regression Coefficients for Analysis of VO2max Change from Baseline – AEECS	X	X
Table 15.2.2.1.4: Variable Selection in Analysis of VO2max Change from Baseline – AEECS	X	X
Table 15.2.2.1.5: Descriptive Statistics of VO2max Change from Baseline - AEECS	X	X
Table 15.2.2.1.6: Analysis of VO2max Change from Baseline (Acute Effect, V4) – Full Analysis Set		
Table 15.2.2.1.7: Analysis of VO2max Change from Baseline (Training Effect, V43) – Full Analysis Set		

Table 15.2.2.1.8: Regression Coefficients for Analysis of VO2max Change from Baseline – Full Analysis Set		
Table 15.2.2.1.9: Variable Selection in Analysis of VO2max Change from Baseline – Full Analysis Set		
Table 15.2.2.1.10: Descriptive Statistics of VO2max Change from Baseline – Full Analysis Set		
Table 15.2.2.1.11: Analysis of VO2max Change from Baseline (Acute Effect, V4) – AEBECS		
Table 15.2.2.1.12: Analysis of VO2max Change from Baseline (Training Effect, V43) – AEBECS		
Table 15.2.2.1.13: Regression Coefficients for Analysis of VO2max Change from Baseline – AEBECS		
Table 15.2.2.1.14: Variable Selection in Analysis of VO2max Change from Baseline – AEBECS		
Table 15.2.2.1.15: Descriptive Statistics of VO2max Change from Baseline – AEBECS		
Table 15.2.2.2.1: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Acute Effect, V4) – AEECS	X	X
Table 15.2.2.2.2: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Training Effect, V43) – AEECS	X	X
Table 15.2.2.2.3: Regression Coefficients for Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEECS	X	X
Table 15.2.2.2.4: Variable Selection in Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEECS	X	X
Table 15.2.2.2.5: Descriptive Statistics for Time to Complete the Exercise Capacity Test: Acute and Training Effect Variables – AEECS	X	X
Table 15.2.2.2.6: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Acute Effect, V4) – Full Analysis Set		
Table 15.2.2.2.7: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Training Effect, V43) – Full Analysis Set		
Table 15.2.2.2.8: Regression Coefficients for Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – Full Analysis Set		
Table 15.2.2.2.9: Variable Selection in Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – Full Analysis Set		
Table 15.2.2.2.10: Descriptive Statistics for Time to Complete the Exercise Capacity Test: Acute and Training Effect Variables – Full Analysis Set		
Table 15.2.2.2.11: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Acute Effect, V4) – AEBECS		
Table 15.2.2.2.12: Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Training Effect, V43) – AEBECS		
Table 15.2.2.2.13: Regression Coefficients for Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEBECS		

Table 15.2.2.2.14: Variable Selection in Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEBECS		
Table 15.2.2.2.15: Descriptive Statistics for Time to Complete the Exercise Capacity Test: Acute and Training Effect Variables – AEBECS		
Table 15.2.2.2.16: Sensitivity Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Acute Effect, V4) – AEECS		
Table 15.2.2.2.17: Sensitivity Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline (Training Effect, V43) – AEECS		
Table 15.2.2.2.18: Regression Coefficients for Sensitivity Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEECS		
Table 15.2.2.2.19: Variable Selection in Sensitivity Analysis of Change in Time to Complete the Exercise Capacity Test from Baseline – AEECS		
Table 15.2.2.2.20: Sensitivity Analysis of Descriptive Statistics for Time to Complete the Exercise Capacity Test: Acute and Training Effect Variables – AEECS		
Table 15.2.2.3.1: Analysis of the Intensity of the Exercise Training (Change from Baseline in Cumulative Work Produced) at V14, V28 and V42– AEECS		X
Table 15.2.2.3.2: Regression Coefficients for Analysis of the Intensity of Exercise Training (Change from Baseline in Cumulative Work Produced) – AEECS		
Table 15.2.2.3.3: Descriptive Statistics for Intensity of Exercise Training - AEECS		X
Table 15.2.2.3.4: Descriptive Statistics for Intensity of Exercise Training by Intervals at V5, V14, V29 and V41 - AEECS		X
Table 15.2.2.4.1.1: Analysis of the Changes in Blood Composition Parameters from Baseline (Training Effect, V43) – AEECS		X
Table 15.2.2.4.1.2: Regression Coefficients for Changes in Blood Composition Parameters from Baseline – AEECS		X
Table 15.2.2.4.1.3: Descriptive Statistics of Blood Composition Parameters – AEECS		X
Table 15.2.2.4.2.1: Analysis of the Changes in Respiratory Parameters from Baseline (Acute Effect, V4) – AEECS		
Table 15.2.2.4.2.2: Analysis of the Changes in Respiratory Parameters from Baseline (Training Effect, V43) – AEECS		
Table 15.2.2.4.2.3: Regression Coefficients for Changes in Respiratory Parameters from Baseline – AEECS		
Table 15.2.2.4.2.4: Descriptive Statistics of Respiratory Parameters – AEECS		
Table 15.2.2.4.3.1: Analysis of the Change in Workload from Baseline at a Lactate Level of 4mmol (Exercise Effect, V43) – AEECS		X
Table 15.2.2.4.4.1: Regression Coefficients for Analysis of Heart Rate – AEECS		
Table 15.2.2.4.5.1: Regression Coefficients for Analysis of Oxygen Uptake – AEECS		

Table 15.2.2.4.6.1: Descriptive Statistics of Physiological and Respiratory Parameters by Workload - AEECS		
Table 15.2.2.4.7.1: Descriptive Statistics of RPC (continuous) by Sex – AEECS		
Table 15.2.2.5.1.: Analysis of the Average Step Count by Month – AEECS		X
Table 15.2.2.5.2.: Regression Coefficients for Analysis of Average Step Count by Month – AEECS		
Table 15.2.2.5.3: Descriptive Statistics of mHealth variables - AEECS		X
Table 15.2.2.6.1: Analysis of the Changes in Biological Health Markers from Baseline at V14, V28 and V43 – AEECS		
Table 15.2.2.6.2: Regression Coefficients for Analysis of Changes in Biological Health Markers from Baseline – AEECS		
Table 15.2.2.6.3: Descriptive Statistics of Biological Health Markers - AEECS		
Table 15.2.2.6.4: Analysis of the Changes in Biological Health Markers from Baseline at V14, V28 and V43 – AEBECS		
Table 15.2.2.6.5: Regression Coefficients for Analysis of Changes in Biological Health Markers from Baseline – AEBECS		
Table 15.2.2.6.6: Descriptive Statistics of Biological Health Markers - AEBECS		
Table 15.2.2.7.1: Analysis of the Changes in Body Fat and Waist Circumference from Baseline (Acute Effect, V4) – AEECS		
Table 15.2.2.7.2: Analysis of the Changes in Body Fat and Waist Circumference from Baseline (Exercise Effect, V43) – AEECS		
Table 15.2.2.7.3: Regression Coefficients for Analysis of the Changes in Body Fat and Waist Circumference from Baseline– AEECS		
Table 15.2.2.7.4: Summary of Body Fat and Waist Circumference by Product Use Category – AEECS		
Table 15.2.2.7.5: Analysis of the Changes in Body Weight from Baseline by Month – AEECS		
Table 15.2.2.7.6: Regression Coefficients for Analysis of the Changes in Body Weight from Baseline– AEECS		
Table 15.2.2.7.7: Summary of Body Weight by Product Use Category – AEECS		
Table 15.2.2.8.1: Descriptive Statistics for Exhaled CO – AEECS		
Table 15.2.2.8.2: Descriptive Statistics for Biomarkers of Exposure - AEECS		X
Table 15.2.2.8.3: Descriptive Statistics for Exhaled CO – AEBECS		
Table 15.2.2.8.4: Descriptive Statistics for Biomarkers of Exposure - AEBECS		
Table 15.2.3.1.1: Overview of Adverse Events – Safety Set		X
Table 15.2.3.1.2: Summary of Adverse Events by System Organ Class and Preferred Term – Safety Set		X
Table 15.2.3.1.3: Summary of Adverse Events Related to Investigational Product and Expectedness, System Organ Class and Preferred Term - Safety Set		X
Table 15.2.3.1.4: Summary of Adverse Events by Severity, System Organ Class and Preferred Term - Safety Set		

Table 15.2.3.1.5: Summary of Adverse Events Leading to Product Discontinuation, Interruption or Reduction by System Organ Class and Preferred Term – Safety Set		X
Table 15.2.3.1.6: Summary of Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term – Safety Set		X
Table 15.2.3.1.7: Summary of Serious Adverse Events by System Organ Class and Preferred Term – Safety Set		
Table 15.2.3.1.8: Summary of Adverse Events related to Study Procedure by System Organ Class and Preferred Term – Safety Set		
Table 15.2.3.1.9: Summary of Frequent Adverse Events (Incidence >5%) by System Organ Class and Preferred Term – Safety Set		
Table 15.2.3.1.11: Summary of Device Events and Malfunctions - Safety Set		X
Table 15.2.3.1.12: Summary of Medical History by System Organ Class and Preferred Term – Safety Set		
Table 15.2.3.1.13: Concomitant Medications by ATC and Preferred Drug Name – Safety Set		X
Table 15.2.3.1.14: Concomitant Diseases by System Organ Class and Preferred Term – Safety Set		
Table 15.2.3.1.15: Prior Medications by ATC and Preferred Drug Name – Safety Set		X
Table 15.2.3.2.1: Summary of Hematology Parameters by Product Use Category – Safety Set		
Table 15.2.3.2.2: Summary of Clinical Chemistry Parameters by Product Use Category – Safety Set		
Table 15.2.3.2.3: Summary of Urinalysis Parameters by Product Use Category – Safety Set		
Table 15.2.3.4.1: Summary of Vital Signs by Product Use Category – Safety Set		X
Table 15.2.3.4.2: Descriptive Statistics of ECG – Safety Set		X
Table 15.2.3.4.3: Descriptive Statistics of Spirometry at Screening, Discharge or Early Termination – Safety Set		X
Table 15.2.3.4.4: Summary of Physical Examination – Safety Set		X
Table 15.2.3.4.5: Summary of Cough-VAS and Likert Scales – Safety Set		X
15.3 Listings		
Listing 15.3.1.1: Discontinued subjects – Safety Set		
Listing 15.3.1.2: Study Enrollment – Screened Subjects		
Listing 15.3.1.3: Subject disposition - Screened Subjects		
Listing 15.3.2.1: Protocol deviation – Safety Set		
Listing 15.3.3.1: Subjects excluded from the As Exposed Exercise Compliant Set – Safety Set		
Listing 15.3.4.1: Demographic data – Safety Set		

Listing 15.3.4.2: Product Preference – Screened Subjects		
Listing 15.3.4.3: Violation of inclusion- and exclusion criteria – Screened Subjects		
Listing 15.3.4.4: Medical history and concomitant diseases – Safety Set		
Listing 15.3.4.5: FTND Questionnaire – Safety Set		
Listing 15.3.4.6: IPAQ Questionnaire – Safety Set		
Listing 15.3.4.7: Lifestyle Questionnaire – Safety Set		
Listing 15.3.5.1: Smoking history – Safety Set		
Listing 15.3.5.2: Product Use Diary – Safety Set		
Listing 15.3.6.1: Body weight and body fat – Safety Set		
Listing 15.3.6.2: Bike Training – Safety Set		
Listing 15.3.6.3: VO2 Max Test – Safety Set		
Listing 15.3.6.4: Rate of Perceived Capacity Scale – Safety Set		
Listing 15.3.6.5: Exercise Capacity test – Safety Set		
Listing 15.3.6.6: Blood Composition and Volume – Safety Set		
Listing 15.3.6.7: CO Breath test – Safety Set		
Listing 15.3.6.8: Cough assessment – Safety Set		
Listing 15.3.6.9: Borg, Lactate levels, Heart Rate and Oxygen Levels during VO2max Test		
Listing 15.3.7.1: Pre-treatment adverse events – Safety Set		
Listing 15.3.7.2: Listing of Adverse Events – Safety Set		
Listing 15.3.7.3: Listing of Deaths and Other Serious Adverse Events – Safety Set		
Listing 15.3.8.1: Safety analysis - Clinical chemistry – Safety Set – Parameter [Unit]		
Listing 15.3.8.2: Safety analysis - Serology – Safety Set – Parameter [Unit]		
Listing 15.3.8.3: Safety analysis - Urinalysis – Safety Set – Parameter [Unit]		
Listing 15.3.8.4: Biomarker – Blood Sample - Safety Set – Parameter [Unit]		
Listing 15.3.8.5: Biomarker – Urine Spot Sample– Safety Set – Parameter [Unit]		
Listing 15.3.9.1: Vital signs – Safety Set		
Listing 15.3.9.2: Physical Examination – Safety Set		
Listing 15.3.9.3: ECG – Safety Set		
Listing 15.3.9.4.1: Spirometry – Safety Set		
Listing 15.3.9.4.2: PRE Broncholidator Spirometry test – Safety Set		
Listing 15.3.9.4.3: Spirometry: Salbutamol/albuteriol inhalation – Safety Set		
Listing 15.3.9.4.4: POST Bronchodilator Spirometry test – Safety Set		
Listing 15.3.9.5: Urine Pregnancy test – Safety Set		
Listing 15.3.9.6: Alcohol Breath Test – Safety Set		
Listing 15.3.9.7: Urine Drug Screen – Safety Set		

Listing 15.3.9.8: Urine Cotinine Test – Safety Set		
Listing 15.3.10.1: Concomitant procedures – Safety Set		
Listing 15.3.10.2: Device events – Safety Set		
Listing 15.3.10.3: Device Accountability – Safety Set		
Listing 15.3.10.4: Qualitative Interview – Safety Set		
Listing 15.3.10.5: Visit Dates – Safety Set		
Listing 15.3.10.6: Comments – Safety Set		
Listing 15.3.10.7: Vivofit Activity – Safety Set		
15.4 Statistical Output Listings		
Table 15.4.2.1.1: Analysis of VO2max Change from Baseline – AEECS		
Table 15.4.2.2.1: Analysis of Change in Exercise Capacity from Baseline – AEECS		

Document Information

Document Name	RD_SMF01_810308_SAP_P1-EXC-01-EU
Document Title	Excap SAP version 2
Version	1.0
Lifecycle	Approved Lifecycle
Status	Approved

Signature Table

Approver	Date (UTC time)	Reason for signing	Outcome
[redacted]	12-Mar-2020 15:03:22	Author Approval	Approved
[redacted]	12-Mar-2020 15:03:57	Subject Matter Expert Approval	Approved
[redacted]	12-Mar-2020 15:07:11	Business Owner Approval	Approved
[redacted]	12-Mar-2020 20:16:03	Subject Matter Expert Approval	Approved