Non-Interventional Study
Integrated Analysis Plan

Integrated Analysis Plan

Protocol Identification No.	MS202559_0001
Title	A study to investigate the association of real-world sensor-derived biometric data with clinical parameters and patient-reported disease status for monitoring disease activity in COPD patients
Investigational Medicinal Product(s)	Not applicable. No medical device or medicinal product is investigated in this study.
***Integrated* Analysis Plan Author**	**Coordinating Author** CMS, Merck — PPD [redacted]
	Function — **Author(s) Name / Role** PPD [redacted] — [redacted]
***Integrated* Analysis Plan Date and Version**	10 November 2022 / Version 1.0 FINAL
***Integrated* Analysis Plan Reviewers**	**Function** — **Name / Role** Bioelectronics, Merck — PPD [redacted] CMS, Merck — PPD [redacted]

AUTHORS / SIGNATURE PAGE

Merck / External (Palleos healthcare) Responsible	Author	Responsibility	Signature	Date
Written by	PPD	Main author		
Written by		Co-author, modeling tasks		
Written by		Co-author, modeling tasks		
Approved by				
Reviewed / Amended by				

MODIFICATION HISTORY

Unique Identifier for Version	Author	Date of iAP Version	Changes from the Previous Version
1.0	PPD	2022-11-10	Initial version

LIST OF ABBREVIATIONS OF TERMS

Term	Definition
ADE	Adverse device effect
AE	Adverse event
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
CAT	COPD Assessment Test
CI	Confidence interval
COPD	Chronic Obstructive Pulmonary Disease
EC	Exclusion criteria
FU	Follow-up
iAP	Integrated Analysis Plan
IC	Inclusion criteria
ICH	International Committee on Harmonization
ID	Identifier
INN	International Proprietary Name
IQR	Interquartile range
IR	Incidence rate
MI	Multiple Imputation
PN	Preferred Name (WHODrug dictionary structure component)
ROC	Receiver Operator Curve
SOP	Standard Operating Procedure
TLF	Tables, Listings and Figures

Integrated [redacted] 1

1 [redacted] 6

1 [redacted] 6

2 [redacted] 7

3 [redacted] 8

3 [redacted]

3 [redacted] 9

2 [redacted] 0

2 [redacted] 0

2 [redacted]

2 [redacted] 0

2 [redacted]

3 [redacted]

2 [redacted] 4

3 [redacted] 4

3 [redacted]

2 [redacted] C [redacted]

3 [redacted] 7

3 [redacted] 7

3 [redacted] 7

3 [redacted] 7

3 [redacted]

3 [redacted] 8

3 [redacted] 8

3 [redacted]

3 [redacted]

4 [redacted] 2

5 [redacted] 3

6 [redacted] 9

1 INTRODUCTION

This document presents the Integrated Analysis Plan (IAP) for the study MS202559_0001.

The IAP is written to reflect the content of the study protocol (v2.0, 2022-10-21).

The objective of the IAP is to describe the data analysis methodology and the way the study results will be presented for the final analysis that informs the clinical investigation report.

1.1 OBJECTIVES AND OUTCOMES

	Objective	Outcome	IAP section
OBJECTIVES (PRIMARY)	1. To describe the distribution of sensor collected data on physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern, by severity of COPD exacerbations.	Primary Outcome: COPD exacerbations Sensor data ○ physical activity ○ (resting) heart rate ○ heart rate variability ○ temperature ○ respiration rate ○ ratio of in- vs. expiration time ○ cough frequency ○ sleep pattern	3.4.4.1
	2. To predict moderate or severe COPD exacerbations by building a statistical model employing sensor-derived data and demographic and medical covariates.		3.4.4.2
OBJECTIVES (SECONDARY)	3. To evaluate the correlation of sensor collected data with a) patients' self-assessed health status and symptoms (CAT) at baseline and study end, b) lung function and lab values at baseline and study end, c) and with history of mild, moderate, and severe exacerbations at study end.	Secondary Outcome: Health status by CAT questionnaire summary score Lung function by plethysmography Laboratory values Number of COPD exacerbations per type	3.4.4.3
	4. To investigate whether changes in physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern in the 3-month	Secondary Outcome: Health status by CAT questionnaire summary score	3.4.4.4

	observation period, are associated with changes of the patient self-assessed symptoms (CAT).		
	5. To predict the CAT score and changes thereof by building a statistical model employing sensor-derived data and demographic and medical covariates.	Health status by CAT questionnaire summary score	3.4.4.5

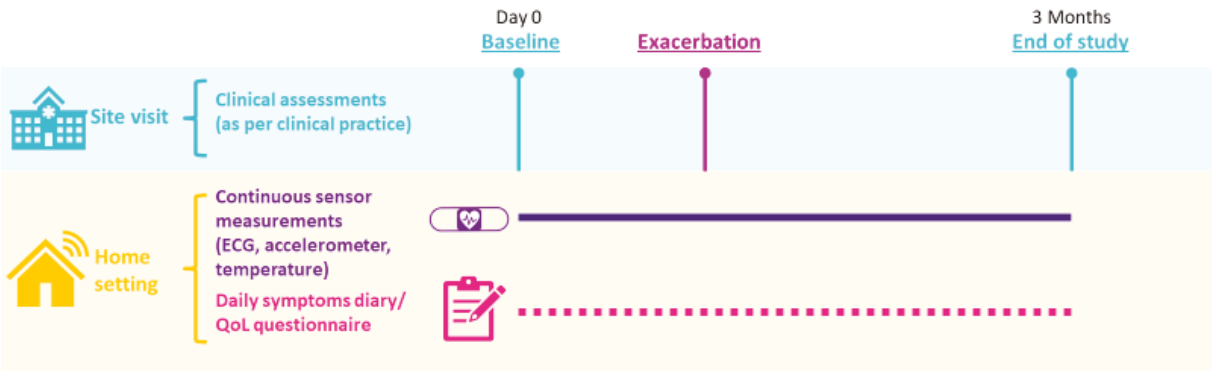
[redacted]

[redacted] STUDY TIME FRAME

COPD Cohort

Participants will be enrolled consecutively, during regular scheduled clinical visits. The Baseline visit (Day 0) marks the index date and can take place once the participant has provided written informed consent to participate in the study. Enrolled COPD participants will be followed for a 3-month observation period (until End of Study visit) or until premature discontinuation from study, whichever occurs first. A look back period of 3 years is used to determine prior history of COPD exacerbations.

Figure 1 Study design overview



Participants in the calibration Cohort

Participants in the calibration cohort will be enrolled in an internal medicine center in Germany. The Baseline visit (Day 0) can take place once the participants provided written consent to participate in the study and marks the index date. Participants in the calibration cohort will be followed for a 2-week observation period.

1.3 STUDY POPULATION

1.3.1 Study Population, Eligibility criteria, Inclusion Criteria

The Study population consists of a COPD cohort and a calibration cohort. The COPD cohort comprises participants with COPD stage II-IV and a corresponding medical history of up to 3 years. Participants in the calibration cohort will be enrolled in an internal medicine center in Germany.

The study population will be identified according to the below inclusion and exclusion criteria. For inclusion in the study, all of the following inclusion criteria must be fulfilled:

For COPD participants:

1. Participants capable of giving signed informed consent, as described in Section 10.1 of the protocol, which includes compliance with the requirements and restrictions listed in the ICF and this protocol
2. Participants ≥40 and ≤80 years at baseline
3. Diagnosis of COPD stage II to IV (documented post-bronchodilator FEV1/FVC < 0.7 and FEV1 < 80% predicted)
4. History of moderate or severe exacerbations (≥2 moderate exacerbations or ≥1 severe exacerbations in any 12-month time window during last 3 years prior to inclusion, considering that last 12 months may reflect lower exacerbation rate due to Covid-19 measures)

For participants in the calibration cohort:

1. Participants capable of giving signed informed consent to participate, as described in Section 10.1 of the protocol, which includes compliance with the requirements and restrictions listed in the ICF and this protocol
2.
3. Participants ≥40 and ≤80 years at baseline

Participants are not eligible for this study if they fulfill any of the following exclusion criteria:

For COPD participants:

1. Participants who, at the discretion of the Investigator, are unlikely to complete the study or will not be able to provide reliable information for the study
2. Clinically relevant and/or serious concurrent medical conditions including, but not limited to visual problems, severe mental illness or cognitive impairment, musculoskeletal or movement disorders, cardiac disease (e.g., heart failure, arrythmia [esp. atrial fibrillation and conduction blocks]), lung cancer (currently treated) that in the opinion of the Investigator, would interfere with participant's ability to participate in the study or draw meaningful conclusions from the study
3. Participants with a cardiac pacemaker, defibrillators, or other implanted electronic devices
4. Participants with known allergies or sensitivity to silicon or hydrogel
5. Less than 6 weeks since previous moderate/severe exacerbation

For participants in the calibration cohort:

1. Participants who, at the discretion of the Investigator, are unlikely to complete the study or will not be able to provide reliable information for the study
2. Clinically relevant and/or serious concurrent medical conditions including, but not limited to visual problems, severe mental illness or cognitive impairment, musculo-skeletal or movement disorders, cardiac disease (e.g., heart failure, arrythmia [esp. atrial fibrillation and conduction blocks]), lung cancer (currently treated) that in the opinion of the Investigator, would interfere with participant's ability to participate in the study or draw meaningful conclusions from the study
3. Participants with a cardiac pacemaker, defibrillators, or other implanted electronic devices
4. Participants with known allergies or sensitivity to silicon or hydrogel
5. Diagnosis of pulmonary disease including, but not limited to COPD, asthma, pulmonary fibrosis, with impact on the lung function and exercise capacity

1.3.2 Study size

Sample size calculation is based on the second part of the primary objective, especially on predicting exacerbations of a severe or moderate type as outcome.

Assuming a desirable sensitivity and specificity of 0.85 for our algorithm and requiring that the lower 95% confidence limit should be at least 0.50 (with a probability of 0.95) for this exploratory approach, the number of observed exacerbation events needed is 18 (1).

Assuming that COPD patients within our eligibility criteria (GOLD stage II-IV with a history of at least 2 moderate to severe exacerbations within any 12 month period within the last 3 years and at least 1 exacerbation within the last 12 months) show a frequency of 2 moderate or severe exacerbations per year, COPD participants will have 0.5 events in 3 months (the in-study time). Further assuming that event occurrence over 3 months for every individual participant follows a Poisson distribution with parameter λ =0.5, the number of events in the overall study population is also Poisson distributed and the corresponding parameter is the sum of individual parameters λ. Calculations show that a λ of 30 is needed to reach an upper cumulative distribution probability of 0.99 for observing at least 18 exacerbations. Therefore, 60 participants are needed to observe at least 18 exacerbation events within the 3-month observation period.

The events of interest in the context of the present study are COPD exacerbations (severe or moderate), exactly the onset of an exacerbation. For building the binary classifier predictive models we also require “no exacerbation” events. We will sample such events randomly from periods with no exacerbations, avoiding 2 weeks before and after an exacerbation.

Considering a drop-out rate of 10%, the estimated number is 67 COPD participants that need to be included in this study.

This recruitment number is also in good agreement with published studies with regards to the ability to build predictive models for exacerbations. Studies with similar recruitment numbers (2) have been able to predict exacerbations based on single parameters studied. Strong features can compensate for low numbers in COPD machine learning model building as shown in (3), and using several features improves the accuracy of predictive models as seen in Jensen 2012.

The study will also observe 10 participants in a calibration cohort for a limited time (2 weeks). This calibration cohort is intended to obtain sensor parameter ranges for individuals who are not suffering from COPD or other pulmonary diseases. These parameter ranges will be used to calculate distance measures of parameters (e.g., physical activity) for COPD patients, and obtain additional sensitivity measures for the sensors used with regards to distinguishing COPD symptoms from a calibration population. The number of participants for these secondary objectives is chosen on an exploratory basis, no sample size calculation has been conducted here. Since we expect no major perturbations (e.g., exacerbations) within the calibration population, we deem a time period of 2 weeks sufficient for generating a representative sample of the participants sensor data.

2 VARIABLES

2.1 VARIABLE DEFINITIONS

2.1.1 Sociodemographic

- Demographics data at baseline: sex, age at Baseline visit, height, and weight, race and ethnicity, age at COPD diagnosis (date of COPD diagnosis - date of birth / 365.25)

- Body mass index (BMI [kg/m^2]) will be calculated based on height and weight:
    - BMI = weight [kg] / (height [cm]/100)^2
- Smoking status/history: cigarette usage (never, current, former), cigar usage (never, current, former), pipe usage (never, current, former), e-cigarette usage (never, current, former), water pipe usage (never, current, former), chewing tobacco (never, current, former); *frequency per day/week/month/year will also be considered within each category*

2.1.2 Medical/Medication history, comorbidities, vital signs

- Medical history at baseline
- Comorbidities at baseline
- Current medication including COPD treatment at baseline
- Date of diagnosis of COPD
- COPD severity (GOLD stage) at baseline
- Number of moderate or severe exacerbations prior to baseline visit (preceding 36 months, in case available)
- Physical examination/vital signs
    - Systolic Blood Pressure
    - Diastolic Blood Pressure
    - Pulse
    - Temperature
    - Respiration Rate

2.1.3 Lung function and laboratory values

The following variables will be measured at baseline, at each exacerbation visit, and at end of study visit.

- Current or last available body plethysmography parameters:
    - FEV1 [L],
    - FVC [L] (forced vital capacity),
    - FEV1/FVC [],
    - TLC [L] (total lung capacity),
    - RV [L] (residual volume),
    - FRCpleth [L] (functional residual capacity),
    - IC [L] (inspiratory capacity),

  - Raw [kPa s L-1] (=ratio of sRaw to FRCpleth; airway resistance),
  - sRaw [kPa s] (specific airway resistance),
  - TV [L] (mean tidal volume),
  - ERV [L] (expiratory reserve volume),
  - IRV [L] (inspiratory reserve volume)
- Blood gas analysis:
  - (capillary) pO2 [mmHg]
  - O2 saturation [%]
  - pCO2 [mmHg]
- Serum CRP [mg/l] and serum procalcitonin levels [ng/ml]
- Blood count (CBC with differential):
  - Hb [g/dL]
  - Hct [%]
  - Erythrocyte count [10^6/µl]
  - MCV [fL]
  - MCH [pg]
  - MCHC [g/dL]
  - Reticulocytes [x10^3/µL] [‰ (of erythrocytes)]
  - RDW []
  - platelet count [10^3/µL]
  - WBC count [x10^3/µL]
  - Lymphocytes [x10^3/µL] [%]
  - Monocytes [x10^3/µL] [%]
  - Neutrophils [x10^3/µL] [%]
  - Eosinophils [x10^3/µL] [%]
  - Basophils [x10^3/µL] [%]

2.1.4 Sensor measurements by Vivalink wearable device and ePRO diary

Unless indicated otherwise, the following variables are measured continuously throughout the observation period.

- Physical activity including stair climbing, (calculated based on continuous accelerometer measurement) [daily average] (derived parameter)

- Heart rate as beats per minute [daily average]
- Resting heart rate (resting state as determined by accelerometer, beats per minute) [daily average] (derived parameter)
- Heart rate variability reflecting differences in time intervals between 2 R-waves in the ECG (milliseconds) [daily average] of the following parameters:
    - SDRR (standard deviation of RR intervals),
    - SDNN (standard deviation of NN intervals),
    - SDNNI (Mean of the standard deviations of all the NN intervals for each 5 min segment of a 24 h HRV recording),
    - RMSSD (Root Mean Sum of Squared Distance),
    - ln(RMSSD) (natural logarithm of RMSSD),
    - pNN50 (percentage of successive RR intervals that differ by more than 50 ms)
    - SI (stress index),
    - LF power (low frequency fraction of HRV, 0.04 – 0.15 Hz),
    - HF power (high frequency portion of HRV, 0.15 – 0.40 Hz),
    - LF/HF (ratio of LF-to-HF power) (derived parameter)
- Temperature [daily average, minimum and maximum values] (derived parameter)
- Respiration rate (breath per minute) [daily average, minimum and maximum values]
- Ratio of in- vs expiration time [daily average, minimum and maximum values] (derived parameter)
- Cough frequency (derived parameter, count per day)
- Sleep pattern (derived parameter, hours of sleep)
- Self-reported activities as collected by the participant app, logged as activity per time-stamp, allowing for several logs of the same activity per day:
    - coughing,
    - coughing with sputum production,
    - spray use,
    - non-scheduled spray use,
    - eating,
    - going to bed,
    - waking up,
    - walking,
    - climbing stairs,
    - and physical exercise

2.2 FOLLOW-UP TIME

- Enrolled COPD participants will be followed for a 3-month observation period (index date until End of Study visit) or until premature discontinuation from study, whichever occurs first.
- Participants in the calibration cohort will be followed for a 2-week observation period from index date.

2.3 OUTCOME DEFINITIONS

2.3.1 Outcomes for primary objectives

The primary outcomes are based on COPD exacerbations. Exacerbations are classified as

1) mild if they are treated with short-acting bronchodilators only,
2) moderate if they are treated additionally with antibiotics or oral corticosteroids, or
3) severe if the patient visits the emergency room or requires hospitalization because of an exacerbation.

The occurrence of exacerbations during the observation period is measured as follows.:

- For moderate/severe exacerbation requiring a physician's intervention, exacerbation details will be recorded via eCRF by the clinical site. Severity, onset and duration will be derived according to the definition above from the documented COPD medication associated with the patients' exacerbation visit. The definition details for COPD medication and ATC codes will be finalized during the planned Data review meeting prior to the final analysis.
- For mild exacerbations: self-reported; treated by the participant without intervention by a physician will be documented in the patient application labeling section (days with added treatment: short-acting bronchodilators only). The start date and end date of addition of short-acting bronchodilators can thus be retrieved from the patient application. The documented use of additional medication (Bedarfsmedikament) will be considered as a mild exacerbation with the following rules for start and end days:
    - A mild exacerbation must start with at least 2 consecutive days of additional medication reported, the first day of consecutive medication will determine the start of the exacerbation.
    - If there are gaps in the additional medication reporting, a maximum of 2 days is allowed to be considered only one exacerbation. A minimum of 70% of the days must have additional medication reported. The end date of the exacerbation is the last day where both these rules are fulfilled.

- Any exacerbations that happen within 5 days will be considered as overlapping, as well as exacerbations that have an intersection in their start and end dates. Overlapping exacerbations will be considered as only one and classified as the graver one. (4)
    - If all overlapping exacerbations have the same type, the start date is the start date of the first exacerbation, and the end date of the last one.

- If the overlapping exacerbations have different types, the start date will be the start date of the first moderate or severe exacerbation, and the end date is the end date of the last moderate or severe exacerbation.

The following exacerbation-related outcomes are required for the analysis of primary objectives:

1. Time points (start and end dates) of mild/moderate/severe exacerbations for descriptive analysis of sensor variables (see Section 2.1.4)
2. Occurrence of moderate/severe exacerbations as binary outcome measured at the onset date of an exacerbation for predictive analysis. An event is the occurrence of a moderate/severe exacerbation. Non-exacerbation events (for definition see Section 1.3.2) will be sampled randomly from time points (days) outside of exacerbation periods (defined by start and end dates of each exacerbation) and avoiding 2 weeks before the onset time (day) and two weeks after the end time (day) of an exacerbation.

2.3.2 Outcomes for secondary and CCI objectives

Outcomes to Achieve Secondary Objectives

Secondary objective	Outcome definition	Number of objective
To evaluate the correlation of sensor collected data with a) patients´ self-assessed health status and symptoms (CAT) at baseline and study end, b) lung function and lab values at baseline and study end, c) and with history of mild, moderate and severe exacerbations at study end.	a) Patients' health status and symptoms at baseline (Day 0) and study end will be measured as the summary score across items of the CAT questionnaire that consists of 8-items in which participants can choose a score from 0 to 5, for each visit b) Lung function will be assessed using plethysmography and lab values including Complete Blood Count with differential, Blood Gas Analysis, procalcitonin and CRP will be assessed as per standard practice at baseline (Day 0) and at study end (for details see Section 2.1.3) c) number, date of onset and duration of mild, moderate, and severe exacerbations, respectively, during the observation period	3
Sensor derived data may be related to the CAT score, therefore, an additional objective is to investigate whether changes in physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern in the 3-month observation period, are associated with changes	Please see above (3a) During the observation period the CAT score will be obtained on a daily basis. Daily summary score across items of the CAT questionnaire will be computed.	4

of the patient self-assessed symptoms (CAT).		
To build a model for predicting patient self-assessed symptoms (CAT) score on daily basis and change in the CAT score (and change of CAT score by 2 points or more considered as MCID) employing sensor collected data on physical activity, heart rate, heart rate variability, temperature, respiration pattern, cough frequency, and/or sleep pattern.	Patients' health status and symptoms at baseline (Day 0) will be measured as the summary score across items of the CAT questionnaire that consists of 8-items in which participants can choose a score from 0 to 5. During the observation period the CAT score will be obtained via a digital application on a daily basis.	5

CCI

CCI

[redacted]

2.3.3 Safety outcomes

Adverse Events

Adverse events are continuously collected in both cohorts throughout the observation period and coded according to current version of WHO MedDRA dictionary.

Adverse device effects (ADE)

Adverse events possibly, probably, or causally related to the use of a medical device or procedures.

This includes any adverse event resulting from insufficient or inadequate instructions for use, deployment, implantation, installation, operation, or any malfunction of the medical device used in the study. This includes any event that is a result of a use error or intentional misuse.

3 DATA ANALYSIS

3.1 OVERVIEW OF PLANNED ANALYSES

There will be a single analysis after last-patient-out during the study close-out phase to support the clinical investigation report.

3.2 ANALYSIS SETS

The study will have two main analysis sets:

- COPD cohort: all patients with COPD fulfilling all of the inclusion criteria and none of the exclusion criteria (see Section 1.3.1);
- Calibration cohort: participants in the calibration cohort as defined in Section 1.3.1

3.3 AMENDMENTS TO THE STATISTICAL ANALYSIS PLAN

Not applicable.

3.4 STATISTICAL ANALYSIS

Standard analyses and descriptive statistics will be carried out using R 4.2 or higher (The R Foundation for Statistical Computing, Vienna, Austria). Linear and mixed models will be fit using SAS 9.4 or higher (SAS Institute Inc., Cary, NC, USA) or using R 4.2 or higher (The R Foundation for Statistical Computing, Vienna, Austria). Machine learning models will be implemented using Python 3.8 or higher (The Python Software Foundation, Wilmington, Delaware, USA).

The creation of tables, listings and figures will follow the descriptions provided in Appendix 1) TLF plan.

3.4.1 General specifications

Descriptive statistics will be provided for study variables and measurement time points as specified in the following sections. Unless indicated otherwise, descriptive statistics will include the number of observations and missing values, means, standard deviations, medians, interquartile range, Q1, Q3, minimum, and maximum. Results for continuous variables (e.g., age at diagnosis) and correlation coefficients will be reported with at least 3 figures as rounded numbers with 2 decimal values. Descriptive statistics for categorical variables will report n (%), missing and with number (n) equal to the total number of patients at each category level (without missing) during the assessment period and will be reported with 3 figures as percent with 1 decimal value.

Medical history, adverse events and prior/concomitant medication will be summarized using coded preferred terms by proportion of patients with at least one occurrence relative to the number of patients in the cohort.

3.4.2 Considerations for potential confounders and effect modifiers

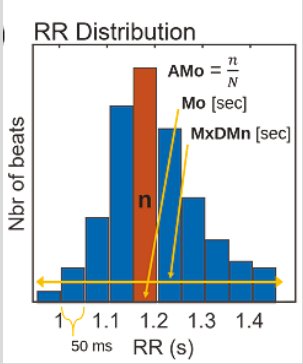
Not applicable. All available features will be considered in the predictive modelling.

For objective 4, smoking status, and other relevant baseline covariates will be forced into the models.

3.4.3 Derived variables

Variable	Original data	Description of derived data	Algorithm to develop/Detail for implementation
Body mass index (BMI [kg/m^2])	Measured height and weight. (Demographics data at baseline)	Weight (kg) / (Height [cm]/100)^2	NA

Physical activity, including stair climbing	Continuous accelerometer measurement, self-reported activities	Classification algorithm, similar to the one described in (5)	An activity flag is extracted from the accelerometer by Vivalink, by using a predefined threshold for adult movement. For stair climbing, first periodic movement must be determined, by using frequency analysis on specific time windows, and generating a ratio to the total spectrum indicating periodic activity over a certain threshold (See (5) for more details). Climbing of stairs is determined by a classification algorithm, here a tree classifier and a nearest neighbors classifier will be tested, using the activity labels inputted by the patient for supervised learning.
Heart rate	ECG	Beats per minute.	Heart rate is provided by Vivalink. In case it needs to be calculated (e.g., for initial comparison of results obtained), the process to follow is to use the heartpy python package, by using the 'process' method. See (6) for details on how to use the package. Said method returns two datasets, one of them being the measures, where one can extract the beats per minute (bpm). *working_data, measures = heartpy.process(ecg_data, ecg_fs)* *Heart_Rate= measures['bpm']* Where ecg_fs is the sample rate of the ECG, 128 in this case.
Resting heart rate	Accelerometer, ECG	Resting state as determined by accelerometer, Beats per minute extracted from ECG	The resting heart rate is calculated as described in the heart rate calculation,but is calculated only in periods where the activity flag is 0.
Heart rate variability reflecting differences in time intervals between 2 R-waves in the ECG (milliseconds) SDRR, SDNN, SDNNI, RMSSD, ln(RMSSD), pNN50, SI , LF power, HF power, LF/HF	ECG	For details of calculation see (7)	ECG peaks are calculated as explained for the heart rate calculation. The inter beat time (inter peak) is measured - RR- and the measures derived from it are: SDRR, the standard deviation of intervals between heartbeats SDNN, the standard deviation of intervals between heartbeats, after removing abnormal beats $$SDNN = \sqrt{\frac{1}{X-1}\sum_{i=1}^{X}(NN_i - \overline{NN})^2}$$ SDNNI, mean of the standard deviations of all the NN intervals for each 5 min segment of a 24-h HRV recording RMSSD, the root mean square of successive differences between adjacent N-N intervals $$RMSSD = \sqrt{\frac{1}{X-1}\sum_{i=1}^{X}(NNdiff_i)^2}$$ ln(RMSSD), natural logarithm of RMSDD

			pNN50, percentage of adjacent NN intervals that differ from each other by more than 50 ms SI, the Baevsky's stress index, calculated as AMo / 2Mo x MxDMn. Where AMo is the amplitude of the mode presented in percent in comparison to all other RR intervals. Mo is the mode of the RR intervals. MxDMn is the difference between the maximum and minimum measured RR interval.  (8) Applying a Fast Fourier Transformation (FFT) or autoregressive (AR) modeling one can separate HRV into its component ULF, VLF, LF, and HF rhythms that operate within different frequency ranges. Given in absolute values of power (milliseconds squared). LF power, low frequency power (0.04–0.15 Hz) HF power, high frequency power (0.15–0.40 Hz) LF/HF Ratio, spectral HRV index computed as (LF/HF)
Tempera-ture	Display temperature	Relative values	Temperature is provided by Vivalink. The value for temperature is derived by Vivalink from the display temperature and then calibrated using initial calibration values, in an IP protected process. The sensor temperature is considered only as a relative value to evaluate changes in the temperature, and not as an objective human body temperature value, meaning no thresholds relative to normal human body temperature are considered, and it will not be used as a marker for fever or hypothermia.
Respiration rate	ECG	Measure of R-R peaks, and application of algorithms as the ones described in (9)	Respiration rate is provided by Vivalink. In case it needs to be calculated, the process to follow is to use the heartpy python package, by using the 'process' method. See (6) for details on how to use the package. Said method returns two datasets, one of them being the measures, where one can extract the breathing rate. *working_data, measures = heartpy.process(ecg_data, ecg_fs)*

			Resp_rate = measures['breathingrate'] Where ecg_fs is the sample rate of the ecg, 128 in this case.
Ratio of in-vs expiration time	ECG	Measure of R-R peaks, and application of algorithms as the ones described in (9)	This measure will be calculated with the heartpy python package, by using the 'process' method. See (6) for details on how to use the package. Said method returns two datasets, one of them being the measures, where one can extract the breathing signal. *working_data, measures = heartpy.process(ecg_data, ecg_fs)* *breathing_signal = working_data['breathing_signal']* Where ecg_fs is the sample rate of the ecg, 128 in this case. Using the breathing signal one can determine the inspiration and expiration peaks (using the peak finder python package scipy.signal.find_peaks, see (10)). The difference between said peaks in milliseconds can be used to determine the ratio of inspiration (distance from lower point to next peak) vs expiration (distance from peak to next lower point). The average, mode, max and min of said ratios will be reported.
Cough frequency	Accelerometer, ECG, self-reported activities	Predictive algorithm similar to the one described in (11)	Accelerometer and ECG signals are filtered to reduce noise from the sensor wear, different filters must be tested to see which one works best for this specific data, including low pass, high pass, bandpass and notch filters. Features are extracted by both dimensionality reduction (PCA, NMF) and frequency analysis, specifically, windows of 2 ot 5 seconds can be taken where measures as mean, median, sd, skewness, kurtosis, etc. are calculated for the first and second component, and the predominant frequency of a window is considered, as well as the magnitude of the signals, and used as features to train a supervised classifier (using the activity labels inputted by the patients). This technique will be tested with random forest, gradient boosting and other boosting methods.

Sleep pattern	Accelerometer, self-reported activities	Self-reported going to bed and waking up, combined with accelerometer data to improve accuracy of sleeping time, plus accelerometer analysis to infer sleep quality.	The basis of the sleep pattern calculations is the self-reported bedtimes. With the same technique as the cough frequency prediction, inactivity signals can be predicted from the labeled data to improve the bedtime accuracy, and the changes in accelerometer (step detection algorithms) can be used to quantify the number of clear breaks in the sleep (standing up, strong cough, etc.). In addition, HR and HRV will be analyzed for unusual patterns (e.g., sudden stronger increases in HR could either be indicative for waking up, or from experiencing sleep apnea. HR can also be used to support information from the accelerometer regarding getting up for example. Expected HR changes during REM phases may be considered. HRV would be expected to increase during sleep. Likewise, temperature changes may support the above analyses. ML algorithms (semi-supervised / self-learning) for classifying periods of sleep disturbances may be explored after building an initial label set using observer consent on periods of sleep disturbance. The result will be the number of hours sleeping per day, a qualifier between 1 and 3 of sleep quality using the average movement during self-reported sleeping hours, being 1 below average movement, 2 average movement, and 3 above average movement and the number and duration of phases of sleep disturbance.
Frequency of additional medication	Self-reported activities	Self-reported use of additional medication	Count of the number of times the use of additional medication (Bedarfsmedikament) as a log activity is reported per day.

3.4.4 Handling of missing data

For variables obtained only at the baseline visit, missing values will not be replaced.

For the self-reported satisfaction questionnaire items, collected at the end of study visit, missing values will not be replaced. For variables measured daily during the observation period such as sensor-generated data, CAT score, and labelling information, missing data will not be imputed for all descriptive statistical analyses performed. The CAT score and derived sensor variables may be imputed for the purpose of building machine learning models if needed to avoid loss of instances using a suitable imputation method (e.g., LOCF, Forward filling, multiple imputation [MI], recursive imputation, median imputation). Sensitivity analyses will be conducted to assess the impact of imputation.

For variables with values collected at baseline, at the exacerbation visit and at end of study visit, missing values will not be imputed for all descriptive statistical analyses. Lung function parameters and lab values may be imputed for the purpose of building machine learning models if needed to avoid loss of instances by LOCF method if the last (previous) available value is recorded within 2 weeks before the date of the visit of interest. In cases where no value for a variable is available for a participant, imputation strategies across instances may be used (e.g., median imputation or iterative imputation methods). Sensitivity analyses will be conducted to assess the impact of imputation.

3.4.5 Analysis of outcomes

3.4.5.1 Primary Objective 1

Descriptive statistics will be reported in terms of mean (SD) or median (IQR) for continuous variables, as well as counts and frequency (%) for categorical variables of sensor collected data described in section 2.1.4, comprised of

- physical activity,
- heart rate,
- resting heart rate,
- heart rate variability (10 variables),
- temperature,
- respiration rate,
- ratio of in- vs. expiration time
- cough frequency,
- sleep pattern (hours of sleep),
- frequency of additional medication

measured for each patient

1. at baseline (Day 0),
2. in the period Day 8 to Day 14 (intra-individual period average),
3. from 14 days prior to an exacerbation onset, in a 7-day sliding time window, in a 3-days sliding time window, and a 1-day sliding time window,
4. at the onset of an exacerbation (first day of exacerbation period),
5. from the end of an exacerbation to 14 days after, in a 7-day sliding time window, in a 3-days sliding time window, and a 1-day sliding time window,
6. 14, 21, and 28 days after an exacerbation.

The analysis will be repeated for different exacerbation severities:

1. Exacerbations of any (moderate or severe) severity are counted
2. Only severe exacerbations are counted
3. Only moderate exacerbations are counted
4. Only mild exacerbations are counted

Results will be reported in one table per parameter, with descriptive statistics indexing rows and measurement times indexing columns.

For each continuous parameter, mean values will be plotted across all daily measurement time points during the 3-month observation period, together with 95% CIs based on the normal approximation (t-statistic).

3.4.5.2 Primary Objective 2

To predict moderate or severe COPD exacerbations by building a statistical model employing sensor-derived data and demographic and medical covariates.

To generate several features needed as input to the model a number of algorithms (e.g., self-learning algorithms) for obtaining derived parameters such as cough frequency, physical activities and sleep pattern will be developed first using the label dataset obtained via the app (see Section 3.4.3).

Different machine learning algorithms will be evaluated, and their performance will be assessed by estimating their accuracy, specificity, sensitivity, and area under ROC metrics. Other usual evaluation metrics such as precision, positive predictive value (PPV), negative predictive value (NPV), F1-score will be computed.

This analysis will be performed on the COPD cohort. Internal cross-validation will be tested.

The following models will be used:

- Logistic regression
- Random Forest
- K-nearest neighbors
- Boosting methods (e.g., Gradient boosting machine, GBM)
- Support vector machines (SVM)

Other methods may be tested.

The development of those models will be based on the machine learning algorithms implementation of scikit-learn (12), a python package designed for predictive data analysis. Other packages may be considered if needed (e.g., H2O (13)).

To optimize the performance of the models, feature extraction and selection algorithms will be used.

Explanatory models will be built on top of the selected machine learning algorithm (14), as a way to support rationalization for physicians and regulatory organizations. Depending on the selected ML model selected, either model specific (for logistic regression, decision trees, k-nearest neighbors) or model agnostic methods will be considered (global methods as PDP, ALE, functional decomposition, and local methods as LIME and Shapley values).

3.4.5.3 Secondary Objective 3

Objective 3 will be analyzed on the COPD cohort. For each pair of variables and outcomes (x, y) defined below, Pearson's correlation coefficient will be computed together with a 95% asymptotic confidence interval based on Fisher's Z transform.

In case of low variability in parameter values (e.g., exacerbation count), parameters may be treated as ordinal and analyzed in terms of Kendall's Tau statistic with non-parametric confidence interval (see (15)).

The following variable groups and outcomes will be paired and assessed at the indicated time points:

Group x	Group y	Time points for assessment
Sensor data (see section 2.1.4), daily averages (17 variables): physical activity, heart rate, resting heart rate, heart rate variability, temperature, respiration rate, ratio of in- vs expiration time, cough frequency, sleep pattern (hours of sleep)	CAT summary score: Self-assessed patient's health status by CAT questionnaire.	Baseline (Day 0), End of study visit
	Lung function parameters by plethysmography (see section 2.1.3, 13 parameters)	Baseline (Day 0), End of study visit
	Laboratory values (see section 2.1.3, 20 parameters): Current or last available (capillary) pO2, O2 saturation and pCO2; Serum CRP and serum procalcitonin levels; blood count (CBC with differential)	Baseline (Day 0), End of study visit
	Exacerbation count by mild, moderate or severe exacerbation.	End of study visit

For each pair of groups (x, y) and for each assessment time point, a table with x and y parameter groups indexing rows and columns respectively, will be provided to display coefficient of correlation and 95% CI for all parameter pairs. To provide a visualization, the tables will be plotted as matrix heatmaps displaying the correlation coefficient point estimates.

3.4.5.4 Secondary Objective 4

Sensor derived data may be related to the CAT score, therefore, an additional objective is to investigate whether changes in physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern in the 3-month observation period, are associated with changes of the patient self-assessed symptoms (CAT).

This analysis will be performed on the COPD participants set.

The following analyses will be conducted in this order:

- Linear mixed model with the patient self-assessed symptoms, CAT score, mean value assessed on a 7-day basis as the response variable and the mean values of physical activity, heart rate,

heart rate variability, temperature, respiration rate, cough frequency also assessed in the same timeframe as the independent variables. Individual trajectories will be investigated. A model including the change of each sensor variable, the time (7-day period) as fixed effect will be tested. Interaction terms between the time and each sensor variable will be included in the model and potentially patient-level random intercepts and slopes will also be included in the model. Other covariates such as age, sex, smoking status, medical history will be considered. The same model will be considered with the absolute change of the CAT score and the absolute change of the sensor variables.

$$Y_{ij} = (\alpha + a_i) + (\beta + b_i) X_{ij} + \xi_{ij}$$

where a_i is the random intercept, b_i is the random slope for the j^{th} observation of the individual i

- The analysis above will be repeated on the COPD participants with severe/moderate exacerbations in the time-window considering the 14 days prior the exacerbations. Daily values of the CAT score and of the physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern will be considered in the model.

  The relation between the change of the patient self-assessed symptoms (CAT score) and the changes of the physical activity, heart rate, heart rate variability, temperature, respiration rate, cough frequency, and sleep pattern during the 3-month observation period will be further evaluated using linear spline regression to infer the location of the changepoints.

3.4.5.5 Secondary Objective 5

To predict the CAT score and changes thereof by building a statistical model employing sensor-derived data and demographic and medical covariates

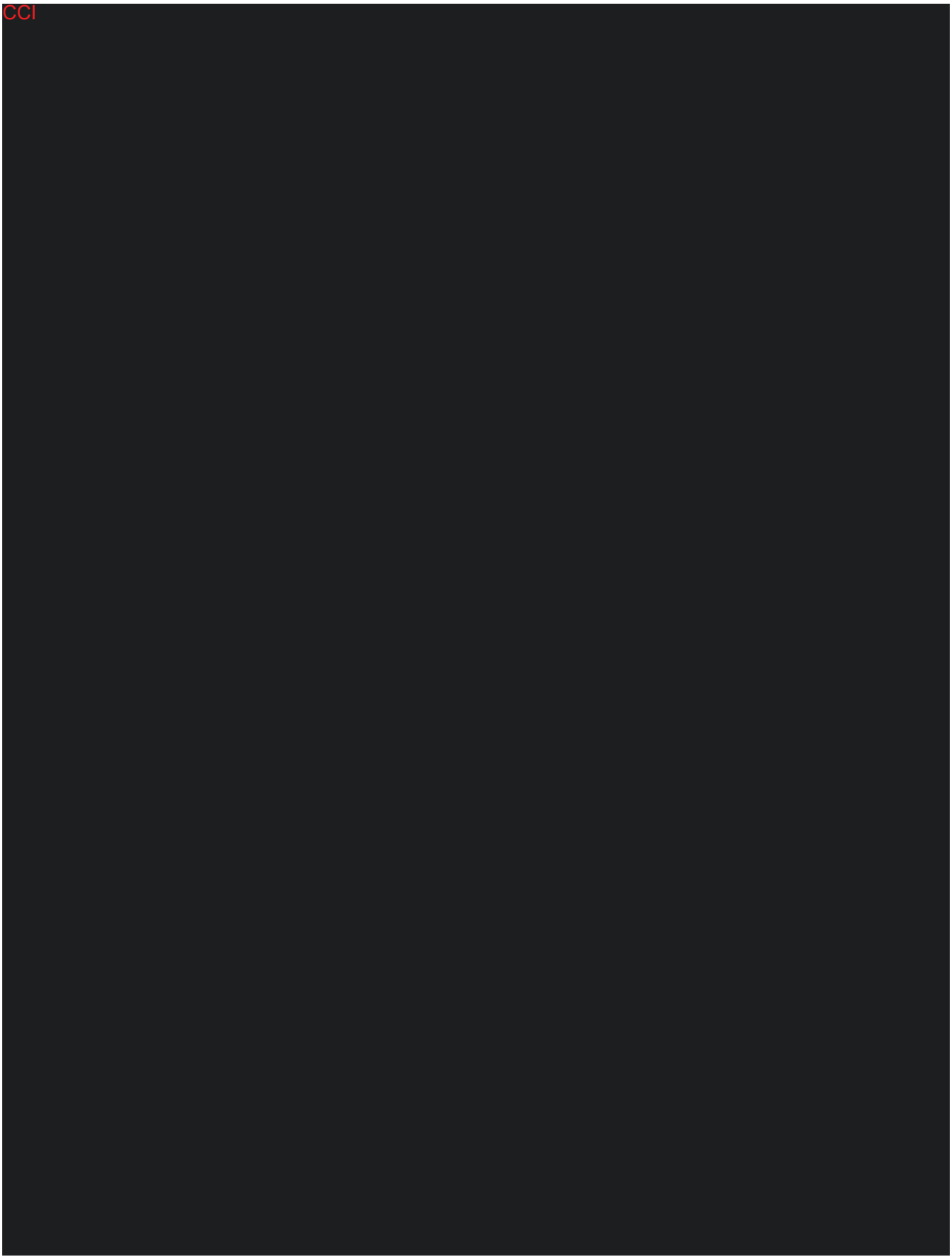
This analysis will be performed on the COPD participants set.

Similar to Primary Objective 2, different machine learning algorithms will be evaluated, and their performance will be assessed by estimating their accuracy, specificity, sensitivity, and area under ROC metrics. Other usual evaluation metrics such as precision, positive predictive value (PPV), negative predictive value (NPV), F1-score will be computed.

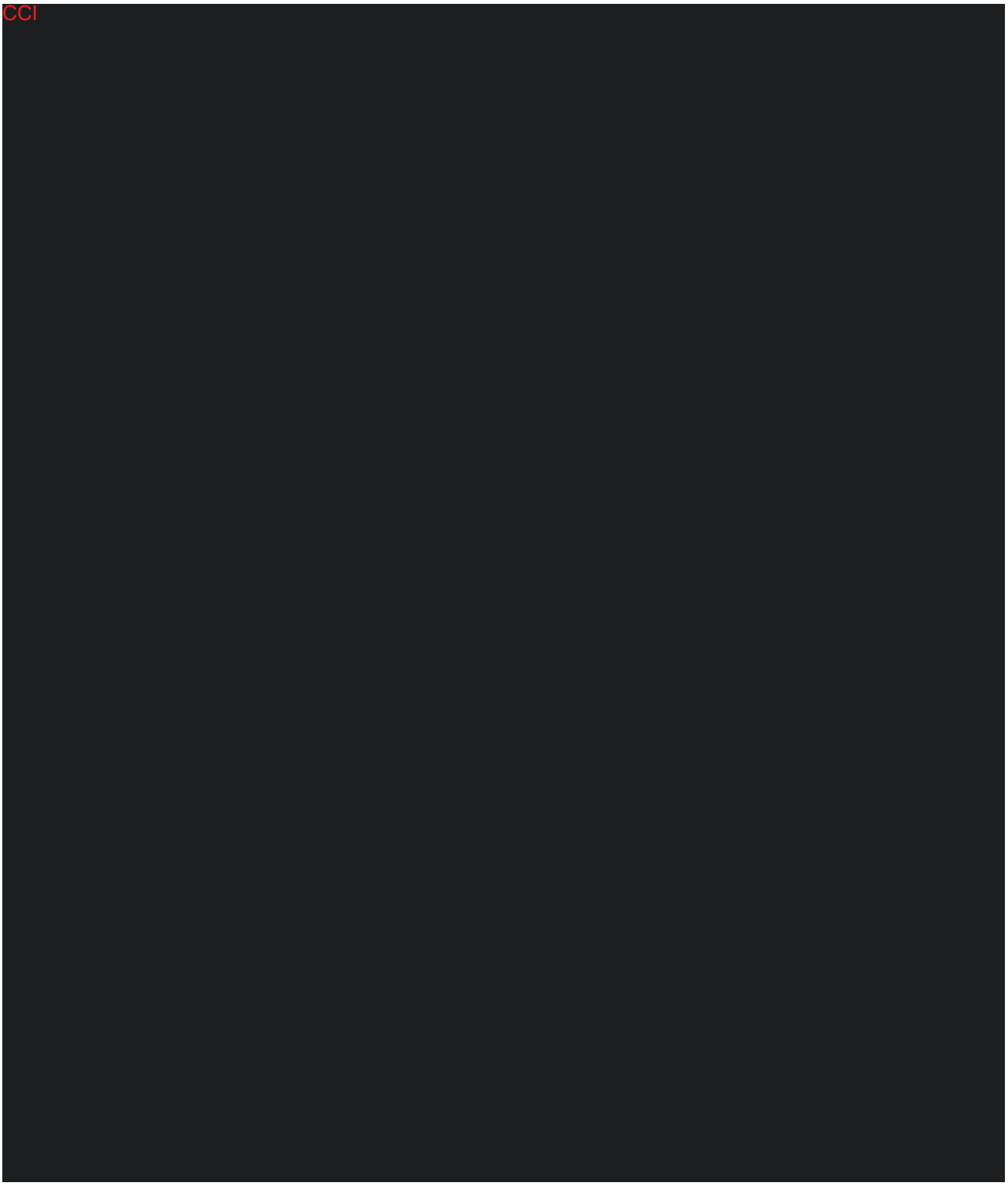
CCI

[redacted]

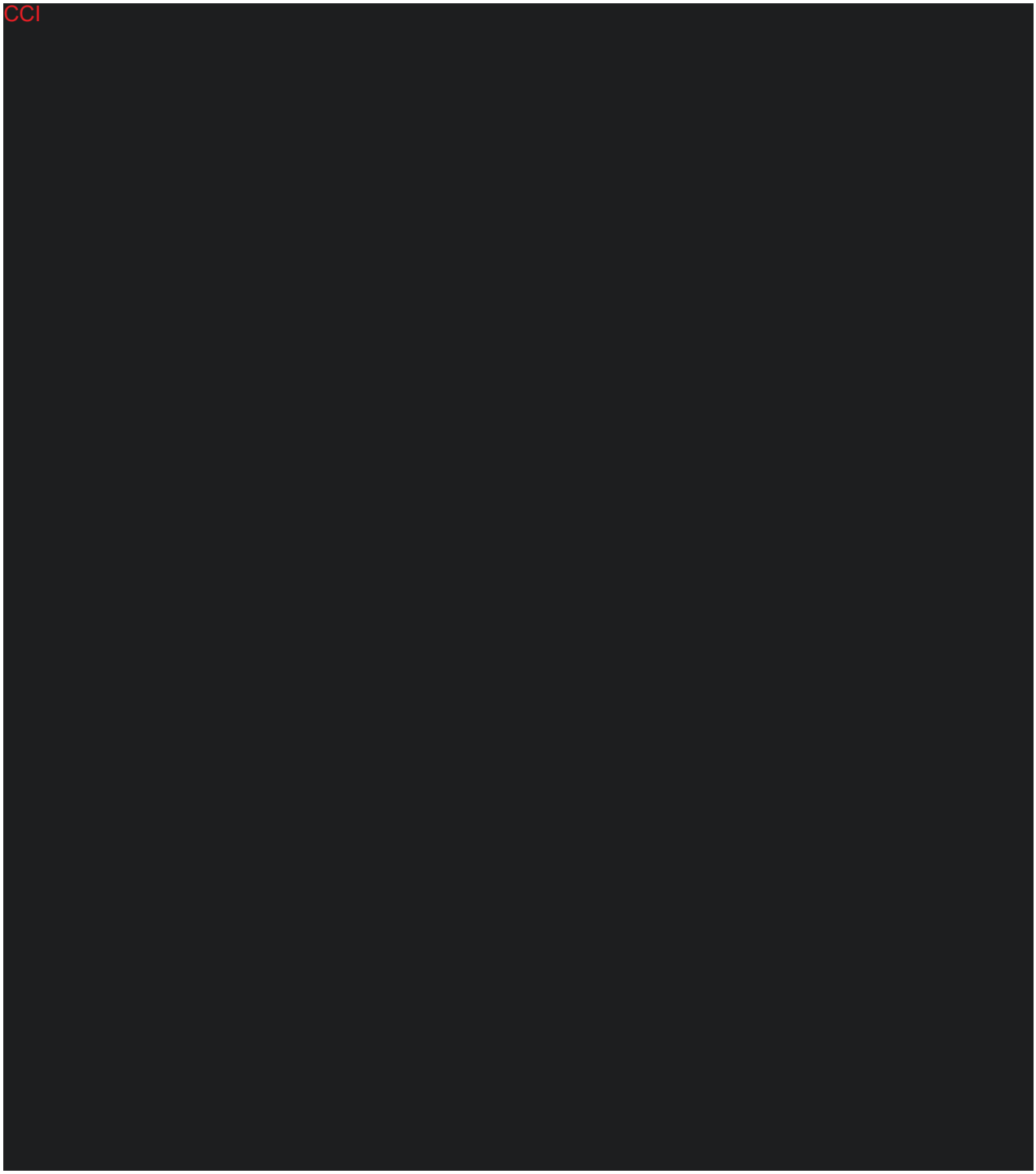
CCI



CCI



CCI



3.4.5.10 Analysis of safety outcomes

Adverse device effects

There is a very low likelihood for ADEs to occur in this study. Any such events will be reported in a listing, containing Subject ID, Site ID, Sex, Age, COPD Severity, Reported Term, Preferred Term, System Organ Class, Serious Flag, Severity, Outcome. ADEs are collected in both cohorts.

Adverse events

Adverse events are continuously collected throughout the observation period in both cohorts. A frequency table will describe AE occurrence by severity.

Proportions of AE will be reported by preferred term and system organ class according to WHO MedDRA dictionary (latest version).

3.4.5.11 Other outcome analyses

COPD exacerbations

For COPD exacerbations, a histogram will be provided showing color-coded occurrence counts of mild, moderate and severe exacerbations by day during the 3-month observation period. In addition, descriptive statistics for exacerbation occurrence and duration will be provided, stratified by severity.

Other clinical outcomes

The following outcome parameter groups (see Sections. 2.1.2, 2.1.3) will be measured on site during exacerbation visits and during the end of study visit for the COPD cohort:

- Physical examination
- Lung function
- Laboratory values

Depending on their overall occurrence, measurement time points will be defined globally as

- Baseline study visit (Day 0),
- First exacerbation visit (Ex 1),
- Second exacerbation visit (Ex 2),
- Ex 3, …, Ex N,
- End of study visit (EOS).

Descriptive statistics will be provided for each measurement. Results will be reported in one table per parameter, with descriptive statistics indexing rows and measurement time points indexing columns.

3.4.6 Sensitivity analyses

To assess effect of missing values, sensitivity analyses will be performed to run models with variables with a missing rate less than or equal to 20% (17).

3.4.7 Subgroup analysis

Not applicable.

3.4.8 Analysis of representativeness

a) Sociodemographic variables will be described in a cross-frequency table, stratified by cohort.
b) Relevant medical history, as well as comorbidities, will be coded by the most recent WHO MedDRA dictionary version to report proportions by preferred term, stratified by cohort, in a cross-frequency table. Rows will be indexed by system organ class and preferred term.
c) Prior and concomitant medication will be coded according to WHODrug Global B3 dictionary, latest version. Proportions will be reported in a cross-frequency table by Preferred Name (sorted according to ATC-levels 2, 4 and 5), stratified by cohort.

COPD cohort only:

d) GOLD stage distribution will be described in a frequency table. The number of moderate and severe exacerbations prior to baseline visit will be described in a frequency table, sorted by documented counts.
e) Smoking status will be described in a frequency table. Smoking history will be reported in a frequency table, stratified by modality (cigarette, pipe, etc.).
f) Lung function will be described by body plethysmography parameters at baseline (reported as part of 3.4.4.11).
g) Laboratory parameters will be described at baseline measurement (reported as part of 3.4.4.11).
h) Vital signs at baseline measurement (reported as part of 3.4.4.11).

3.4.9 Patient disposition

a) The number of patients screened versus enrolled will be displayed, reporting screening failures in absolute values and percent, stratified by reason (inclusion/exclusion criteria violation, consent withdrawn, other) for both cohorts. The number and percentage of patients completing the 3-month (2-week for calibration cohort) follow-up period will be reported, as well as patients with early discontinuation, stratified by reason (consent withdrawn, lost to follow-up, died). Results will be reported in a CONSORT-type diagram.
b) COPD cohort: Relative to the number enrolled, patients with 0, 1, 2 or more exacerbation visits will be described in a frequency table.

3.4.10 Description of Follow-up

Follow-up duration will be described for both cohorts

a) For the subgroup of patients with early study discontinuation, follow-up duration will be described.
b) Reason for end of study will be described for both cohorts

4 REFERENCES

1. *Sample size calculation should be performed for design accuracy in diagnostic test studies.* **Flahault, A., Cadilhac, M. und Thomas, G.** Aug 2005, J Clin Epidemiol, Bd. 58(8), S. 859-62.

2. *Monitoring breathing rate at home allows early identification of COPD exacerbations.* **Yañez, AM, et al.** 2012, Chest, Bd. 142, S. 1524-1529.

3. *Once Daily Versus Overnight and Symptom Versus Physiological Monitoring to Detect Exacerbations of Chronic Obstructive Pulmonary Disease: Pilot Randomized Controlled Trial.* **Al Rajeh, A.M., et al.** 2020, JMIR Mhealth Uhealth, S. 8:e17597.

4. *Time course and pattern of COPD exacerbation onsetThorax.* **Aaron, SD, et al.** 2012, Thorax, Bd. 67, S. 238-243.

5. *Differentiating Between Walking and Stair Climbing Using Raw Accelerometry Data.* **Fadel, W.F., et al.** 2019, Stat Biosci, Bd. 11, S. 334-354.

6. **van Gent, P.** Online documentation. *Heartpy.* [Online] https://python-heart-rate-analysis-toolkit.readthedocs.io/en/latest/.

7. *An Overview of Heart Rate Variability Metrics and Norms.* **Shaffer, F. und Ginsberg, J.P.** 2017, Front Public Health, S. 5:258.

8. **Kubios.** HRV analysis Methods. [Online] 2022. https://www.kubios.com/hrv-analysis-methods/.

9. *A Comparative Study of ECG-derived Respiration in Ambulatory Monitoring using the Single-lead ECG.* **Varon, C, et al.** 2020, Sci Rep, S. 10:5704.

10. **Virtanen, P, et al.** Online documentation. *Scipy.* [Online] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html.

11. **Georgescu, T.** *Classification of Coughs using the Wearable RESpeck Monitor.* Edinburgh : MInf Project (Part 1) Report. University of Edinburgh, 2019.

12. *Scikit-learn: Machine Learning in {P}ython.* **Pedregosa, F., et al.** 2011, Journal of Machine Learning Research, Bd. 11, S. 2825-2830.

13. **H2O.ai.** Documentation. [Online] https://docs.h2o.ai/h2o/latest-stable/h2o-docs/downloading.html.

14. **Molnar, C.** *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable.* 2nd. Munich, Germany : Independently Published, 2022.

15. **Myles, Hollander, Wolfe, Douglas A. und Chicken, Eric.** *Nonparametric statistical methods.* s.l. : John Wiley & Sons, 2013.

16. *Consensus and dissention: A measure of ordinal dispersion.* **Tastle, W. J. und Wierman, M. J.** s.l. : International Journal of Approximate Reasoning, 2007.

17. *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* **JA, Sterne, et al.** s.l. : Bmj, 2009.

18. **Online documentation. *heartpy.* [Online] https://python-heart-rate-analysis-toolkit.readthedocs.io/en/latest/.**

5 APPENDIX

5.1 APPENDIX 1 : TLF PLAN