Statistical Analysis Plan

TRIAL FULL TITLE	Evaluating the Non-Inferiority of Airmod to Capnostream™35 on Respiratory Rate Monitoring and User Experience Clinical Evaluation
SAP VERSION	2.0
SAP VERSION DATE	2022/5/1
TRIAL STATISTICIAN	
Protocol Version (SAP associated with)	3.0
TRIAL PRINCIPAL INVESTIGATOR	<Principal investigator>
SAP AUTHOR(s)	Yuan Ren Cheng

Identifiers: NCT05263791

1 SAP Signatures

I give my approval for the attached SAP entitled < Evaluation the Non-Inferiority of Airmod to Capnostream™35 on Respiratory Rate Monitoring and User Experience Clinical Evaluation > dated < 2022/5/1>

Statistician (Author)
Name: Yuan Ren Cheng

Signature: _____

Date: _____

Statistician Reviewer
Name:

Signature: _____

Date: _____

Principal Investigator
Name:

Signature: _____

Date: _____

2 Table of Contents

3 Abbreviations and Definitions

Provide a list of the abbreviations and acronyms used in the Statistical Analysis Plan (SAP) with definitions. All terms will appear in alphabetical order.

AE	Adverse Event
AirRR	Airmod-scored auscultation sound generated from AS-101
CapRR	Machine-scored Capnography (CapRR) generated from Capnostream™35 (K150272, Medtronic)
CDA	Clinical Data Analyst
CRF	Case Report Form
DSMB	Data and Safety Monitoring Board
DUT	Device under test
IP	Investigational Product
ManARR	Manual-scored auscultation sound originated from AS-101
ManCRR	Manual-scored Capnography (ManCRR) originated from Capnostream™35 (K150272, Medtronic)
PACU	Post Anesthesia Care Unit
PADSS	Post Anesthetic Discharge Scoring System
PAR	Post Anesthesia Recovery
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan

Table 1. Abbreviations and Definitions

4 Introduction

4.1 Preface

Airmod is a software application designed to aid healthcare professionals by monitoring a patient's breathing in real time. It can detect inhalation acoustic sound and provides respiratory rates based on the analysis of the acoustic signals of breathing sounds collected by the electronic stethoscope. This is an observational, prospective, multi-center, pivotal study.

4.2 Scope of the analyses

These analyses will evaluate the non-inferiority of Airmod to Capnostream™35 for respiratory rate

monitoring and user experience clinical evaluation, the results will be included in the clinical study report.

4.3 Study Objectives

The objective of this study is to demonstrate that the performance and safety of Airmod in monitoring respiratory rate is non-inferior to Capnostream™35.

Primary Objective:
To evaluate accuracy and performance of respiratory rate (RR) measurement from Airmod compared to Capnostream™35. The primary objective is to establish the non-inferiority.

Secondary Objectives:
1. To assess the accuracy of respiratory rate measurement by Airmod in comparison with manual-scored auscultation sound during the less sensitive period of Capnography on Capnostream™35.
2. To measure the agreement between AirRR and ManCRR.
3. To evaluate the response time of the first breath detection followed by administration of jaw thrust during the apnea period. The 3 secondary objective is to compare the response time of Airmod versus Capnography.
4. To compare the influence of subjects to variated breath rates on respiratory rate monitoring in bpm as measured by Airmod, manual-scored and Capnostream™35.
5. To evaluate the safety and usability of Airmod.

Manual-scored Capnography (ManCRR) originated from Capnostream™35 (K150272, Medtronic)
Machine-scored Capnography (CapRR) generated from Capnostream™35 (K150272, Medtronic)
Manual-scored auscultation sound (ManARR) originated from AS-101
Airmod-scored auscultation sound (AirRR) generated from AS-101

4.4 Endpoints

Primary Endpoint:
The difference between AirRR and ManCRR, which is greater than -3 establishes non-inferiority. The unit in this calculation is bpm and sampling rate is 1 Hz.

Secondary Endpoints:
1. The accuracy of inhalation classification in the segmentation basis is greater than 82.5% at the 95% CI. The performance metrics (F1-socre, MAPE) at event level will be demonstrated and target for F1-score > 0.8, MAPE < 0.5.
2. To measure the agreement of AirRR and ManCRR on concordance correlation coefficient, and on their mean and bias by Bland-Altman analysis.
3. Use paired t-test to compare the response time of Airmod versus Capnography on Capnostream™35 after administration of jaw thrust. In the target of $p<0.05$ indicates significance.
4.1. The difference of ManCRR minus AirRR and ManCRR minus CapRR indicates non-inferiority at margin of 3 bpm (M1) and sampling rate 1Hz.
4.2. The difference of ManCRR minus AirRR and ManCRR minus CapRR indicates non-inferiority at margin of 2 bpm (M2) and sampling rate 1Hz.
5. The absence of unanticipated serious adverse device effect (USADE) during study period and the questionnaire.

[Time Frame:
Non-Anesthesia Stage: Up to 40 patients will be randomly assigned to non-anesthesia stage assessment, the measurement will be evaluated after enrollment with ICF, at the non-operating room (NOR) area. (5~30 minutes)
Procedure Stage: From firstly anesthetic administration to end of procedure. (Depends on procedure 10mins ~ few hours)
Transportation Stage: During transportation of subjects. (1~10 minutes)
Post anesthesia care unit stage : From subjects arrive post anesthesia care unit (PACU) to discharge from PACU (PAR or PADSS score ≥ 9) (At least 15 minutes)]

5 Study Methods

5.1 General Study Design and Plan

This is a prospective, multi-center, pivotal study. The purpose of this study is to evaluate the non-inferiority of Airmod to Capnostream™35 for respiratory rate monitoring. In the total of 300 subjects enrolled at the OR and NOR area in the hospital, the interim analysis will be performed at the enrollment of 150 subjects during study period. (using OBrien-Fleming to determine α=0.00153) The target population is divided into non-anesthesia, and sedative anesthesia groups. The non-anesthesia group is proposed to enroll at least 40 patients in order to indicate normal patient population in the health care facilitate. The sedative and anesthesia groups will enroll up to 260 patients to cover a wide range of health condition, comorbidity and risk of sedative patients.

Subgroup analysis is planning to indicate the induction period, intraoperative period, and PACU period. And the body mass index (BMI) will be stratified by < 30 and ≥ 30 (at least 30 Number) to further examine the high risk patients group. This study will not be randomized and blinded.

Safety will be assessed by the investigator(s) and will be monitored until the end of the study. All Adverse Events (AEs) and Serious Adverse Events (SAEs) occurring during the study period will be followed until the event is considered stable or until full resolution.
The duration of this study is estimated to be 12 months, while approximately 1.5 months for each individual subject to complete all participant visits.

5.2 Inclusion-Exclusion Criteria and General Study Population

(ICH E3;9.3. ICH E9;2.2.1)

Inclusion criteria:
1. Provision of signed and dated informed consent form.
2. Stated willingness to comply with all study procedures and availability for the duration of the study.
3. Male or female with at least 20 years of age.
4. Fit for intravenous general anesthesia (IVG) as assessed by pre-anesthesia evaluation.

Exclusion criteria:
1. Presence of neck pain or injuries.
2. Under the use of high-flow nasal cannula ventilation.
3. Unable to wear Airmod and Capnostream™35 device related accessories at the investigator's discretion.
4. As a vulnerable population, including legal incapacity or evidence that a subject cannot understand the purpose and risks of the study, regardless of authorized representative support.

5. Unwilling or unable to comply fully with study procedures (including non-toleration of the capnography cannula) due to any disease condition which can raise doubt about compliance and influencing the study outcome.

5.3 Randomization and Blinding

(ICH E3; 9.4.3, 9.4.6. ICH E9; 2.3.1, 2.3.2)
This study will not be randomized and blinded.

5.4 Study Assessments

(ICH E3; 9.5.1. ICH E9; 2.2.2)

Procedures	**Screening** [a]	**Treatment** [a]
Visit Number	1	2
Visit Days	-28 ~ 1	1
Allowed visit window (days)		+14
Informed consent	X	
Criteria evaluation	X	
Medical history	X	
Demographic data [b]	X	
Physical examination [c]	X	
Vital signs [d]	X	
Study intervention administration		X [e]
PAR assessment [f]		X [g]
PADSS assessment [f]		X [g]

AE review and evaluation		X

a. Screening and treatment can be the same day.
b. Demographic data include date of birth and gender.
c. Physical examination includes general appearance, weight and height, HEENT (head, eyes, ears, nose, and throat), mouth, skin, neck (including thyroid), lymph nodes, spine, cardiovascular system, respiratory system, gastro-intestinal system, nervous system, musculoskeletal system, blood and blood forming organs, mental status, and other body systems if applicable.
d. Vital signs include body temperature, sitting blood pressure and pulse rate.
e. The administered anesthetic medication and dosage, surgery type, start and stop time of surgery will also be collected.
f. Post Anesthetic Recovery (PAR) and Post Anesthetic Discharge Scoring System (PADSS) will be assessed upon admission and every 15 minutes after transferring to post anesthesia care unit (PACU).
g. Subject will be discharged when PAR or PADSS score ≥ 9, consider as complete the study.

Table 2. Study assessments

6 Sample Size

(ICH E3; 9.7.2. ICH E9; 3.5)

By referencing predicate device clinical trials on respiratory rate monitoring in the hospital setting, most of the cases recruited 120+ patients. In order to cover a wide range of health conditions, comorbidity, and risk, we proposed to enroll at least 40 non-anesthesia patients and up to 260 sedative / anesthesia patients from pre-anesthesia evaluation clinics (in a total of 300).

Location	Preanesthesia Evaluation	Preprocedural Patient Preparation	Sedative/Analgesic medication induction	Intraoperation	Post anesthesia care unit (PACU)
OR	Enrollment	Device installation / Preparation	Trial period[1*]	Trial period	Trial period[2*]
NORA	Enrollment	Device installation / Preparation	Trial period[1*]	Trial period	Trial period[2*]
Location	**Preanesthesia Evaluation**	**Non-anesthesia state**			**Eligible to discharge**
Prep room, PACU	Enrollment (40)	Device installation and Trial period for 15~30 minutes			Device installation and Trial period for 15~30 minutes

1* Notice : We included sedative/analgesic medical intended for general anesthesia and not intended for general anesthesia.

2* Notice : We followed the guidance that Patients receiving moderate procedural sedation may continue to be at risk for developing complications after their procedure is completed. Slow drug elimination may contribute to residual sedation and cardiorespiratory depression during the recovery period. Therefore, we proposed to track the recording until patient for fill the discharge requirement.

Fig 1. Enrollment plan of subjects

7 General Analysis Considerations

7.1 Timing of Analyses

In the non-blinded total of 300 subjects enrolled, the interim analysis will be performed at the enrollment of 150 subjects (at least 30 non-anesthesia patients enrolled) during the study period. at

the nominal significance level using OBrien-Fleming to determine α=0.00153. The final analysis will be performed at the end of 300 subjects enrolled. The nominal significance level is using OBrien-Fleming to determine α=0.02450.

7.2 Analysis Populations

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 5.2)

In our study design, both experimental and reference devices are set on all patients enrolled. The analysis is comparing the recording of devices on identical subjects. Therefore, all of the intending to treat patients are included in our study. However, if the patients were at risk of life-threatening events during their operation, the medical crew is obligated to proceed with resuscitation and stop the intervention, in such case, the recording could be terminated and the previous data is still counted in our study.

Population	Definition / Criteria	Analyses Evaluated
Safety	● All subjects who are enrolled in the study. ● This population will be based on the operation the subject actually received.	Safety Adverse events/Adverse devices related events.
Intent To Treat	● Comprise of all subjects with at least 5 minutes recording to evaluate the efficacy. ● This population will be based on the recording of both devices to the subjects enrolled.	Efficacy
Per-Protocol	● The per-protocol population is a subset of the ITT population in this study. ● Subjects who were enrolled in the study and recorded by both devices for more than 5 minutes will be part of the ITT groups and the subjects with a protocol violation that is deemed to affect efficacy assessments in the study period, but their data will be	Per Protocol Population.

	excluded from the period(s) affected by the protocol violation for a PP analysis. (For instance, a patient who was administrated with an advanced airway management tool (LMA) should be considered as a protocol violation in the intraoperative period.)	

Table 3. Analysis populations

7.3 Covariates and Subgroups

(ICH E3; 9.7.1, 11.4.2.1. ICH E9; 5.7)

There are no obvious covariates in the study intervention. However, the risk of anesthesia is the primary concern in our study group demographic. We will record the health history and comorbidity of patients to indicate the underline risk of anesthesia. In addition, the patient with a higher BMI (>30) in the sedative/anesthesia group (at least 30 patients) will be further stratified into a subgroup to represent the patient with higher risk. Besides, at least 40 patients will be assigned to a non-anesthesia stage assessment in order to represent the ordinary patients in the health care facilities.

The recording in the sedative/anesthesia group will divide into three periods, 1. Induction, 2. intraoperative, 3. PACU. All data will be analyzed in the same procedure to indicate non-inferiority of Airmod with EtCO2.

7.3.1 Multi-center Studies

(ICH E3;9.7.1, 11.4.2.4. ICH E9; 3.2)
This study is intended to conduct in 2 sites in Taiwan, with an equal number of patients who will be expected to enroll in all sites but competitive.

7.4 Missing Data

(ICH E3; 9.7.1, 11.4.2.2. ICH E9;5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

The missing data in this study is referring to the non-detecting period of each device, which is excluded from the final comparison. The mean non-detecting period will be reported as a baseline variable in the units of minutes. The non-detecting period is the intrinsic NA value from the device record. It could represent the estimation is out of detecting range, or the device is calibrating itself (which is a basic function of the Capnostream™35). All of the recorded data will be controlled and monitored by Data and Safety Monitoring Board (DSMB) to ensure data integrity.

In our study, the NA value of intrinsic estimation will show an “NA” or “- -” to the user interface of the device, just as the other vital sign monitor. The user interface delivers a clear message for the

malfunction period. It will not affect or increase the risk of respiratory rate measurement.

7.5 Interim Analysis and Data Monitoring

Per FDA required,(ICH E3; 9.7.1, 11.4.2.3. ICH E9; 4.1, FDA Feb 2010 “Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics”). A Data and Safety Management Board (DSMB) will organize the regular meeting by the charter of DSMB. The interim analysis and data/safety monitoring should be evaluated in the meeting with regular meeting minutes. Data will be monitored under standard operation procedure of contract research organization and site management organization to prohibit protocol deviation.

7.5.1 Purpose of Interim Analysis

Due to uncertainty about the efficacy aspect, the interim will allow learning to influence the subsequent design of the study at DSMB. The whole ITT data will be analyzed in the interim analysis to evaluate the objective of this study. By evaluating the efficacy in the primary endpoint at this stage can determine the early success of its non-inferiority.

7.5.2 Planned Schedule of Interim Analysis

The interim analysis will be performed at the enrollment of 150 subjects (must include at least 30 non-anesthesia state patients enrolled) during the study period.

7.5.3 Stopping Rule

We’re proposing not to apply the stopping rule for efficacy analysis in this study. Therefore, a failure end point in the interim analysis will not prohibit the following study. However, by the charter of DSMB. They are obligated to monitor, and terminate the study on a serious protocol deviation, or server medical device related adverse events reported.

7.5.4 Analysis Method to Minimize Bias

The whole ITT data set will be included in the interim analysis. A minimal sample size of 30 patients in each subgroup are required as a constrain to minimize the bias.

7.5.5 Adjustment of p-value

On the confidence interval analysis in this study, we adopt OBrien-Fleming to determine the nominal significance level α=0.00153 at the interim analysis for the 150 samples. Furthermore, at the end of the study the O'Brien-Fleming determines the nominal significance level at 0.02450 to establish the target of our study confidence interval.

7.5.6 Practical Measures to Minimize Bias

- CDA will perform the interim analysis.
- Members of DSMB will see any data or analyses in the interim and make decisions.
- All information will be publically available following an interim analysis.
- All information will be provided to the sponsor and investigators.
- CDA will perform the final analysis.

7.5.7 Documentation of Interim Analysis

Snapshots of the data available at the interim analysis should be preserved, as should all documentation of analysis plans, programming code and reporting provided at the interim. It should be possible to recreate the decision process from the trial archive by CDA.

8 Summary of Study Data

All of the continuous variables in the subject demography will be summarized using the following descriptive statistics: n (sample size), mean, standard deviation. The proportion and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical measures (for instance, gender, principal operations). In general, all data will be listed, sorted by site, and trial period. All summary tables will be structured with a column for each device under test in the order (Airmod, Capnostream™35) and will be annotated with the total population size relevant to that group, including any missing observations.

Only deviations from the general overview will be noted in the subsequent sub-sections within section 9.

8.1 Subject Disposition

All of the subjects are enrolled and completed the study on the same day. There is minimal risk for following up with the candidates.

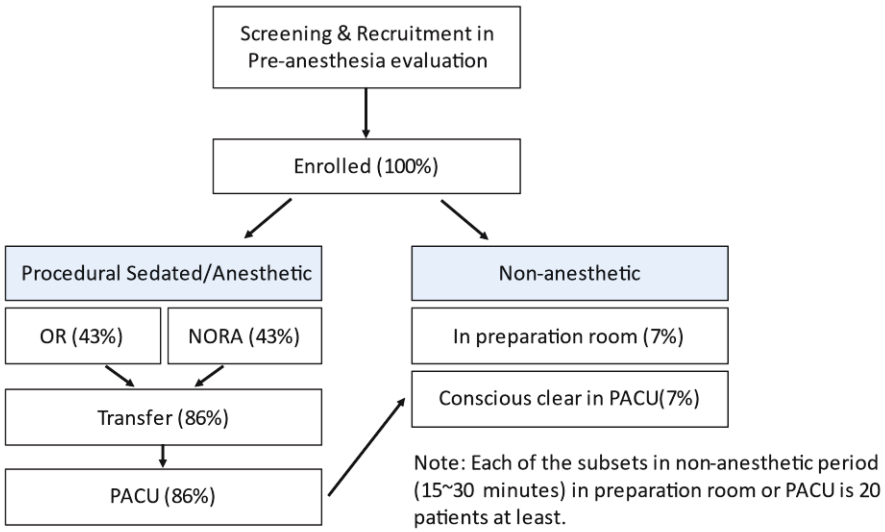


Fig 2. Subjects disposition

The whole screening and recruitment process will take place in the pre-anesthesia evaluation clinics to reveal the detail of this clinical study and we are proposing to enroll 300 patients with their inform-consent form. A subset of patients (20 patients) will be randomly subjected to the non-anesthetic group and record their respiratory rate by both devices in the preparation room. The other candidates will proceed to their original operation/examination and we will record them with both devices through the whole procedural sedation process (from IV induction to discharge of PACU). Within this procedural sedated and anesthetic group a subset of patients (20 patients) will be asked to stay in the PACU for another 15~30 minutes after they are eligible to discharge, which means they are conscious clear and the recording of this period will be allocated to the non-anesthetic group in the future analysis.

The summary statistics will be produced in accordance with section 9.

8.2 Derived variables

8.2.1 Manual-scored Capnography respiratory rate (ManCRR) originated from Capnostream™35 (K150272, Medtronic)

Manual scored capnography respiratory rate is a derived variable from the capnographic recording of Capnostream™35 device, the recording trace will be acquired from the original device record, analyzed by researchers with self developed software, and verified by the principal investigator. The verified data will be further integrated into the case report form as the ManCRR and compared with AirRR.

8.2.2 Manual-scored respiratory rate of auscultation sound (ManARR) originated from AS-101

Manual scored respiratory rate of auscultation sound is a derived variable from the auscultation recording of AS-101 and the original wav file was created by the Airmod. The acoustic waveform will be further processed and displayed on labeling software (Open Source by Heroic Faith medical Science: arXiv:2101.01352). The annotation will be conducted by the researcher in our study and respiratory rate would be estimated upon the analysis of the inhalation and exhalation annotation.

8.2.3 F1-socre

The F1 score is a weighted average of precision and recall for binary classification prediction models. In the secondary end goal “The accuracy of inhalation classification in the segmentation basis” is presented by F1-Score of the ground truth (manual score auscultation sound inhalation event) and the prediction (Airmod inferencing result on inhalation event). We will compare the result on a time series segmentation basis. The confusion matrix is determined by the classification result on manual score auscultation (consider as Actual measurement) and Airmod inferencing result (consider as Prediction measurement).

		Airmod inferencing	
		Inhale	Not Inhale
Manual Scored Auscultation sound	Inhale	**TRUE Positive (TP)**	**FALSE Negative (FN)**
	Not Inhale	**FALSE Positive (FP)**	**TRUE Negative (TN)**

Fig 3. Classification and abbreviation of outcomes of manual scored auscultation sound and Airmod inferencing.

The F1-score is defined as :

Precision = TP / (TP+FP)

Recall = TP/ (TP+FN)

F1-score = 2/ (Precision^{-1} + Recall^{-1})

8.2.4 MAPE

The MAPE (mean absolute percentage error) is a derivative of the accuracy as percentage measurement. The performance is metriced by MAPE of the A_t as the observed value (ManARR) and the F_t as prediction value (AirRR). We will compare the results based on total recording time (by minute).

The MAPE is defined as:

A_t = ManARR

F_t = AirRR

n = Total recording time (by minute)

MAPE = $1/n\ (\Sigma((A_t-F_t)/A_t)$

8.2.5 Bland-Altman analysis

The Bland-Altman analysis is widely used to measure the consistency of continuous data from two distinct measurements. It combines the quantitative analysis of LoA (limits of agreement) and the qualitative analysis of scatter plot to show an instinctive evaluation between two methods of their correlation. Each of the n samples is then represented on the graph by assigning the mean of the two measurements as the x-value, and the difference between the two values as the y-value. The repeatability of RR measurements from Airmod (AirRR) and Capnostream™35(K150272, Medtronic) (CapRR) will be compared with ManCRR separately under the sampling rate: 1 sample point per 60 seconds, the differences are representing into ΔG_A and ΔG_E, while the S(x,y)=((ΔG_A+ΔG_E)/2,ΔG_A-ΔG_E). To compare an estimate of this to the LoA, the two LoA are the mean value of x separately ±1.96SD. (doi: 10.1080/10543400701329422)

8.2.6 Concordance correlation coefficient

The concordance correlation coefficient is a popular indices for assessing the agreement between quantitative measurements taken from different observers (methods). Chia-Cheng Chen et, al. described a repeated measurements method of the concordance correlation coefficient to test the agreement between two measurements (doi:10.1016/j.csda.2012.11.004). To consider that there are N randomly selected subjects where measurements are taken by j observers (fixed) at k time points (random)

Computes Concordance Correlation Coefficient through R language. The epi.ccc of CRAN (Comprehensive R Archive Network) well described the usage and arguments. The formula being used is “epi.ccc(x, y, ci = "asymptotic", conf.level = 0.95, rep.measure = TRUE, subjectid) “, while “x” and “y” representing the sets of two measurements, “ci” indicating the method to be used (z-transform/asymptotic), “conf.level” must be a single number from 0-1, “rep.measure” indicating whether the observations are repeated or not in a logical value (TRUE/FALSE), and “subjectid” is a factor providing details of the observer identifier if rep.measure == TRUE.

Both x and y values need to be present for a measurement pair to be included in the

analysis. If either or both values are missing (i.e. coded NA) then the measurement pair is deleted before analysis.

8.2.7 The outlier of RR estimation

We will use the Tukey approach to identify the outliers of RR estimation, while the difference of AirRR, and ManCRR estimation reach the outside of the interval (Q1-1.5*IQR, Q3+1.5*IQR) will be identified as outliers. The number of outliers in the ITT dataset will be counted and plotted in a box chart on the final report.

8.3 Protocol Deviations

A protocol deviation could take place in circumstances where devices malfunction, calibrations, or interrupting the recording by emergency issues. As aforementioned, missing data of a certain period could be introduced into the final result. The missing data episodes will be reported to the DSMB and eliminated from the final statistical result.

8.4 Demographic and Baseline Variables

Subjects will be demonstrated as the number of patients with percentage in the group. Gender will be recorded into male and female and demonstrated as the number of patients with percentage in the group. Age will be recorded into 5 groups (22-34,35-44,45-54,55-64,65 plus) and demonstrated as the number of patients with percentage in the group. Height, weight, BMI and neck circumference will be recorded, the mean value and standard deviation of them will be calculated and demonstrated. Type of procedures will be recorded, each of the group will be demonstrated as the number of patients with percentage in the group. Duration during procedure and PACU will be recorded, the mean value and standard deviation of them will be calculated and demonstrated. Charles Comorbidity score will be recorded into 16 groups (myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, connective tissue disease, ulcer disease, stroke or transient ischemic stroke, diabetes, hemiplegia, moderate to severe renal disease, diabetes with end organ damage, any tumor, leukemia, lymphoma) and demonstrated as the number of patients with percentage in the group.

8.5 Concurrent Illnesses and Medical Conditions

We will employ the Charlson comorbidity scale and also record the major operation/examination of patients who were enrolled in this study to indicate the concurrent illnesses and medical conditions. Furthermore, the pre-anesthesia evaluation result, including the ASA score, will be also recorded and reported to reveal the patient risk landscape of our study.

The Charlson comorbidity scale records the following disease as reference:

1. Myocardial infarct
2. Congestive heart failure
3. Peripheral vascular disease
4. Cerebrovascular disease

5. Dementia
6. Chronic pulmonary disease
7. Connective tissue disease
8. Ulcer disease
9. Stroke or transient ischemic stroke
10. Diabetes
11. Hemiplegia
12. Moderate to severe renal disease
13. Diabetes with end organ damage
14. Any tumor
15. Leukemia
16. Lymphoma
17. Moderate or severe liver disease
18. Metastatic solid tumor
19. AIDS

Table 4. Diseases recorded in Charlson comorbidity

We will further record and analyze the patient demographic with ICD-10 PCS (procedure codes) to indicate the patient population of concurrent illness during the study.

8.6 Treatment Compliance

Not applicable.

9 Efficacy Analyses

To demonstrate the non-inferiority, the Efficacy Analyses included all of the intent to treat population; it followed the objective of study endpoint to measure respiratory rate performance of Airmod (AirRR) in comparison with the reference device (ManCRR and CapRR). The respiratory rate estimation of each device under test (DUT) is described in the unit of the respiratory count per minute from each measurement over time (sampling rate at 1Hz) regardless of actual cause of inhalation event detected by DUT. The efficacy of agreement in the secondary endpoint is evaluated by Bland Altman analysis, a bias between mean and difference, and 99.69% CI at interim analysis and 95.1% at final analysis of the difference between DUTs in an unit difference plot (unit = breaths per minute). Subgroup analysis of each endpoint focuses on various sections and scenarios of anesthesia (from preparation, induction,

interoperation, transportation, and post anesthesia care unit). An obesity group (BMI >=30) is particularly selected to indicate the high risk patients in the procedural sedation and they will be tested as a stratified subgroup in all of the study endpoints.

9.1 Primary End Point Efficacy Analysis

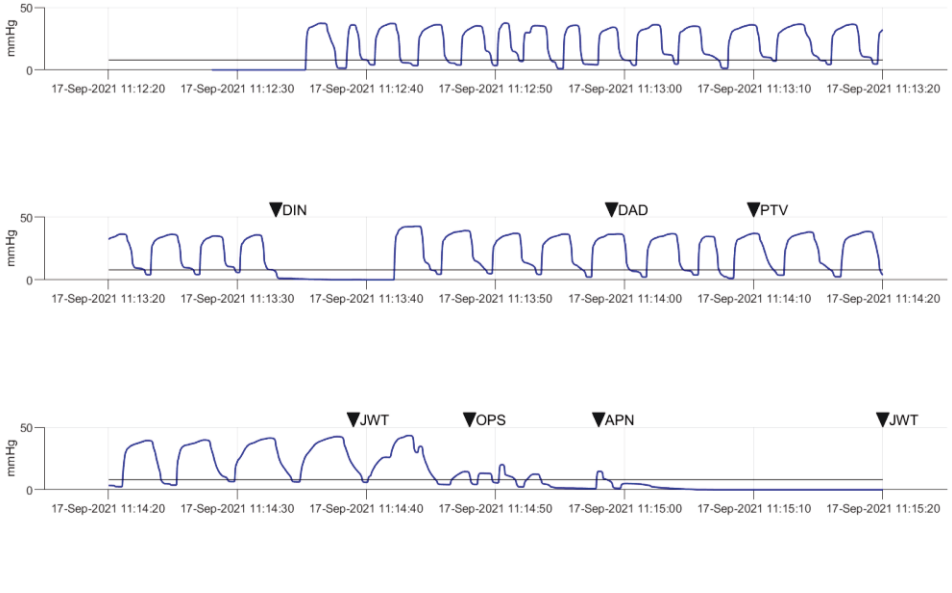
The primary objective to prove the non-inferiority in between the Airmod and reference device is determined by the difference between AirRR and ManCRR, while it is greater than -3 under 99.69% CI at interim analysis and under 95.1% CI at final analysis to establish non-inferiority. The measurement AirRR is a time series list of respiratory rates value (rrUI) recorded by Airmod at 1Hz sampling rate from each subject. The record is retrieved from the Airmod app directly. It also marks the real world time as a time index in the resolution of milli-second.
Sample recording from {type2airm_date_startTime_RR.txt} :

timeindex	rrValue	rrUI
12:37:30.328	23.74	24
12:37:31.137	23.74	24
12:37:32.342	23.74	24

Table 5. Sample recording from {type2airm_date_startTime_RR.txt}

The rrUI out of the claimed range (< 4 Brpm or > 35 Brpm) are excluded prior to efficacy analysis. The case number of the aforementioned record will be presented in the case report form.

The measurement ManCRR is determined by research fellow from their marks of respiratory events on capnography. On the day 1 of the study, the research associate will acquire the capnography data from Medtronic Capnostream™35. The data will subsequently be transcribed to a time series graphical plot, the scale of y axis is identical to the Medtronic Capnostream™35. (Sample plot 1) Research members will mark the respiratory event on the graphical plot with a tailor made software to evaluate the respiratory count per minute at the sampling rate of 1 Hz. A time series list of respiratory rate (ManCRR) will be generated and collected with the original respiratory rate (CapRR) output from the Medtronic *Capnostream™35*. (Sample Recording 1)



Sample plot 1

Fig 4. Capnography data from Medtronic Capnostream™35 transcribed to a time series graphical plot

Sample Recording 1

```
Realtime Full Continuous Transfer
Patient ID      ECKH_20210917_002
Patient Type    Adult
Report Generation Time  Sep 17,21  09:50:50 AM
Date     Time       CO₂ Wave EtCO₂   RR       IPI      SpO₂     PR       A/hr     ODI      EtCO₂ High       EtCO₂ Low        RR High RR Low  No Breath
                    mmHg     mmHg    bpm               %        bpm
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           09:50:50 AM      --       --       --       --       --       --       --       --       0        0        0        0        0        0
Sep 17,21           10:06:06 AM      0.60     27       16       10       98       70       252      252      0        0        0        0        0        0
Sep 17,21           10:06:06 AM      0.61     27       16       10       98       70       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      0.62     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      0.63     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      3.70     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      4.47     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      4.58     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      4.57     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      4.57     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      5.68     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.20     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.36     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.38     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.38     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.37     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.35     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      6.18     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      5.13     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      1.30     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      1.20     27       16       10       98       71       252      252      0        0        0        0        0        0
Sep 17,21           10:06:07 AM      1.20     27       16       10       98       71       252      252      0        0        0        0        0        0
```

Fig 5. Original data (including respiratory rate) output from the Medtronic Capnostream^{TM}35

The real world time index is pre-calibrated in our study, and all of the time series list of respiratory rate will be integrated by time index for efficacy analysis.

The primary efficacy analysis of primary endpoints are comparing the difference of AirRR and ManCRR to reject the null hypothesis (difference of estimation is greater than the non-inferiority limit : -3) under 95% level of confidence interval of all ITT population. The $\mu\alpha$ denotes the respiratory counts of AirRR per minute, and the $\mu\beta$ denotes the respiratory counts of ManCRR per minute. The respiratory count per minute is the mean of recorded samples (eliminating the missing or non-detecting data

point) in each minute. Respiratory counts per minute represent the respiratory rate and are tested in the primary objective hypothesis with the type I error of 0.025, and Type 2 error of 0.2, and the power is 0.8 with single tail non-inferiority analysis.

Reference from our pilot study in 25 cases

Step	Summary
Objective	Non-inferiority: H0:$\mu\alpha-\mu\beta > d_{NI}$ vs H1:$\mu\alpha-\mu\beta \leq d_{NI}$
Endpoint	Margin_1 >-3; Margin_2 >-2
Error	Type I error α=0.025 (one-sided)
	Type II errorβ=0.2, Power 1-β=0.8
Non-inferiority limit	d_{NI}= -3 Brpm
Population standard deviation	σ=3.29647
Others	r=1, dropout rate=0.3%

Table 6. Statistical information referencing from pilot study

9.1.1 Secondary Analyses of Primary Efficacy Endpoint

The primary efficacy analysis will be further divided into three periods of study 1.Induction (3minutes from DIN), 2.Intraoperative period, 3.PACU period to reveal the performance of our device in particular scenarios. The data will be analyzed on a minute basis, and also aggregated per subject basis in each period. The candidates in the non-anesthesia group will be solely analysis for the whole period to represent the normal patient group in a healthcare facility.

Look	Info Fraction	Approx. Cum. Subjects	Nominal Significance Level (1-sided)	Efficacy Boundaries z-scale
Interim	50%	150	0.00153	2.9626
Final	100%	300	0.02450	1.9686

Fig 6. Information of Interim analysis

9.2 Secondary Efficacy Analyses

The secondary endpoints further expand the coverage of respiratory rate measurement scenarios, challenge the M2 margin in the statistical analysis, and try to reveal the responsiveness of airway management in each.

9.2.1 Analysis to assess the accuracy of respiratory rate measurement by Airmod in comparison with manual-scored auscultation sound during the less sensitive period of Capnography on Capnostream™35.

The accuracy of inhalation classification in the segmentation basis is greater than 82.5% at the 95% CI. All of the less sensitive periods shall be pre-determined and validated by the principal investigator.

The 95% CI is based on the number of less sensitive periods. The performance metrics (F1-socre, MAPE) at event level will be demonstrated and target for F1-score > 0.8, MAPE < 0.5.

The accuracy of inhalation classification is based on the classification result of Airmod (Sample recording 2) and manual scored auscultation sound (ManARR) representing actual measurement of the respiratory event (Sample recording 3). The analysis is based on the derived variable F1-score, and MAPE to indicate the target of efficacy.

Sample recording form {type2_date_startTime_AI.txt}:
Sample recording 2

type2_20210917_103934_AI....
檔案(F) 編輯(E) 格式(O) 檢視(V) 說明

```
Label Start End
I 10:39:38.159 10:39:38.734
I 10:39:43.555 10:39:44.053
I 10:39:45.823 10:39:46.436
I 10:39:49.210 10:39:49.746
I 10:39:58.972 10:39:59.508
I 10:40:02.074 10:40:02.495
I 10:40:07.921 10:40:08.304
I 10:40:18.934 10:40:19.546
I 10:40:27.444 10:40:27.865
I 10:40:34.993 10:40:35.759
I 10:40:39.070 10:40:39.797
I 10:40:42.219 10:40:42.717
I 10:40:45.002 10:40:45.729
I 10:40:47.797 10:40:48.487
I 10:40:50.687 10:40:51.453
I 10:40:53.491 10:40:54.104
I 10:40:56.142 10:40:56.908
I 10:41:01.376 10:41:01.912
I 10:41:04.189 10:41:04.840
I 10:41:09.874 10:41:10.448
I 10:41:13.891 10:41:14.312
I 10:42:02.397 10:42:03.048
I 10:42:06.065 10:42:06.640
```

第 27 列 100% Unix (LF) UTF-8

Fig 7. Inhalation classification result output from Airmod

Sample recording 3 (https://doi.org/10.1117/12.2590770)
Noted: < filename of audio_date_research crew.txt>

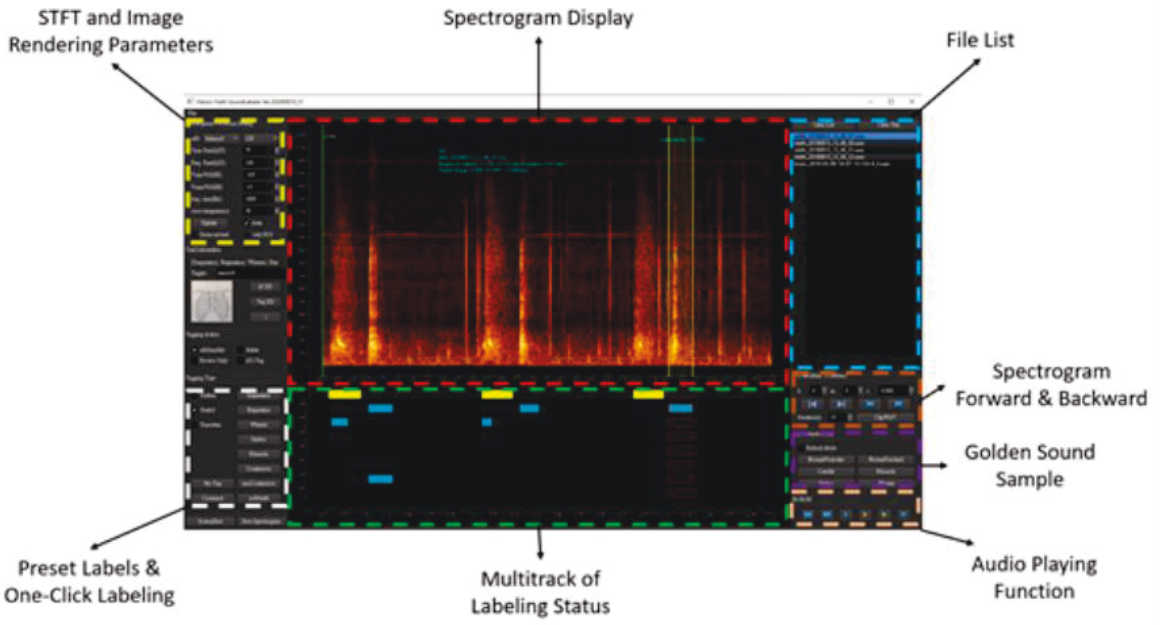


Fig 8. Labeling system sample

The ManARR recording shall be done by at least 2 research crew and the intersection is considered as

the actual measurement of the respiratory event.

9.2.2 Analysis to measure the agreement between AirRR and ManCRR.

Bland Altman analysis, and concordance correlation is adopted to measure the agreement between AirRR and ManCRR. The quantitative results of respiratory rate from these devices under test over time are evaluated by the above mentioned method to analyze the mean and bias from the plot and concordance correlation coefficient.

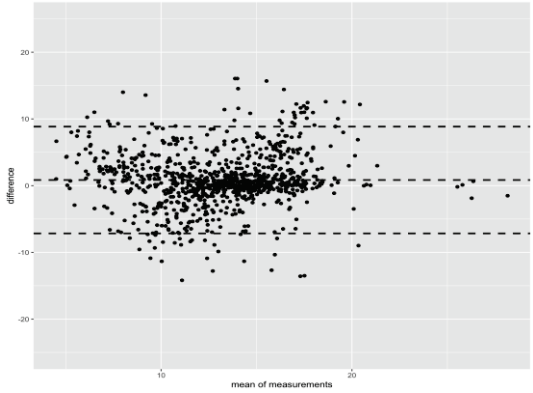


Fig 9. Bland Altman between AirRR and ManCRR (Example)

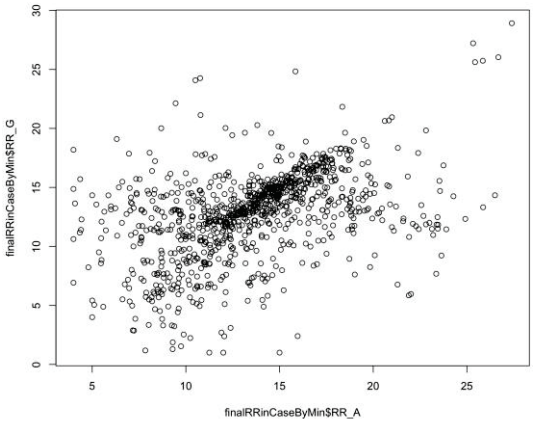


Fig 10. Correlation between AirRR and ManCRR (Example)

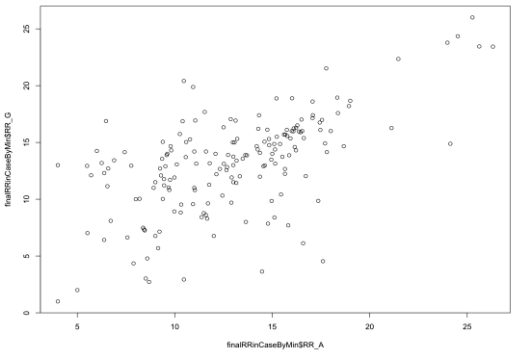


Fig 11. Correlation between AirRR and ManCRR during induction period (Example)

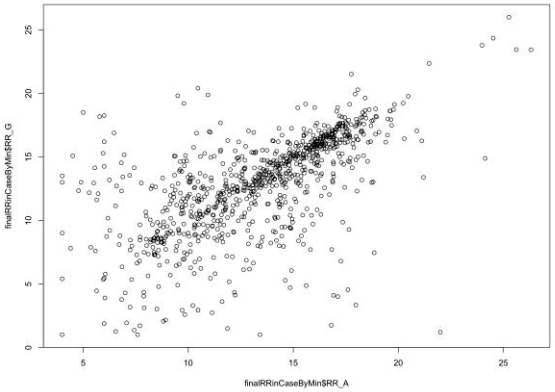


Fig 12. Correlation between AirRR and ManCRR during Inter-operation period (Example)

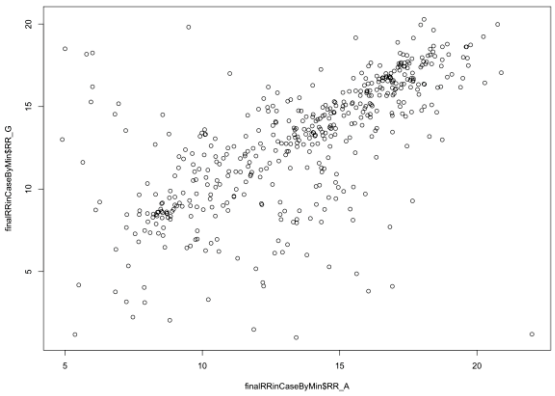


Fig 13. Correlation between AirRR and ManCRR during transportation and PACU period (Example)

Also, we apply a sensitivity test for Airmod by comparing the reference method at each time point and calculating the absolute deviations in breaths per minute (bpm). As a measure of accuracy, absolute deviations were calculated as a percent of reference.

Percent absolute deviation were averaged for each subject yielding overall measures of accuracy. To evaluate the accuracy of Airmod, the individual percent absolute deviation was compared for each monitoring time point by subtracting ManCRR from AirRR and then averaged within subject. A paired t test was used to evaluate significance across subjects.

To measure the precision of Airmod for each subject, the standard deviation of percent deviations (in nonabsolute value) was calculated. The precision of Airmod was compared through a paired t test of the subject-specific standard deviations of Airmod’ s percent deviations from reference.

9.2.3 Analysis to evaluate the response time of the first breath detection followed by administration of jaw thrust during the apnea period. The 3rd secondary objective is to compare the response time of Airmod versus Capnography.

Use paired t-test to compare the response time of Airmod versus Capnography on Capnostream™35 after administration of jaw thrust. The target of $p<0.05$ indicates significance difference.

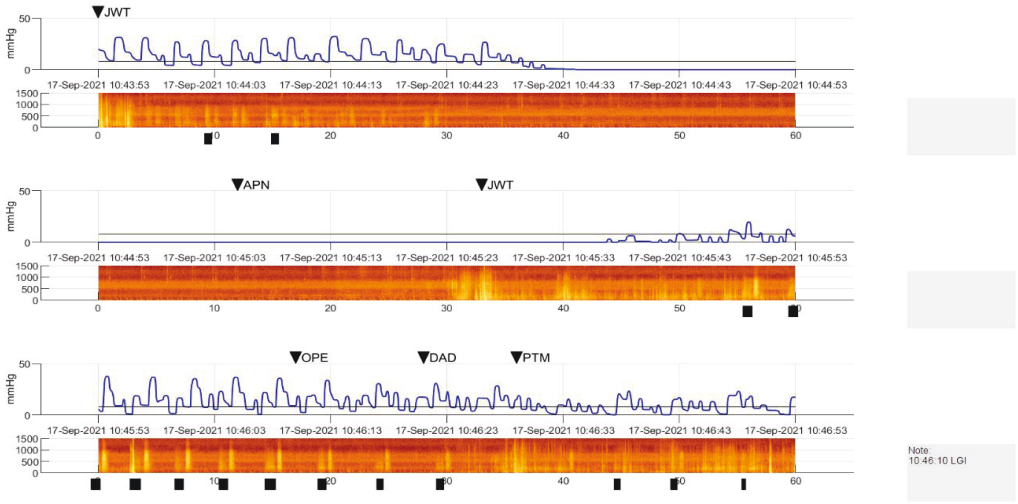


Fig 14. Breath detection (Capnostream™35 & Airmod) followed by administration of jaw thrust during the apnea period

9.2.4 Analysis to compare the influence of subjects to variated breath rates on respiratory rate monitoring in bpm as measured by Airmod, manual-scored and Capnostream™35.

In order to determine the performance of Airmod and Medtronic Capnostream™35 in the ITT population, we compare them with the gold standard (ManCRR). The difference of ManCRR to AirRR and ManCRR to CapRR are parallel listed and compared to indicates non-inferiority at margin of 3 bpm (M1). The range of respiratory rate are also reported as 95% Confidence Interval of measurement. We proposed to further challenge the largest clinical acceptable range. The difference of ManCRR minus AirRR and ManCRR minus CapRR indicates non-inferiority at margin of 2 bpm (M2) are proposed to be the target of efficacy analysis.

9.2.5 Analysis to evaluate the usability of Airmod.

The demography of users will be shown in pie charts in proportion. (Recorded in [CRF_appendix 9_ Questionnaire POC] Part 1)
The evaluation of usability will be calculated from the Part 2 of [CRF_appendix 9_ Questionnaire POC], the pass rate greater than 90% showed that Airmod is good to manipulate.

10 Safety Analyses

All candidates who were subjected to the device with Airmod and Capnostream™35 will be included in the safety analyses (performed in the Safety population) and data will be summarized using descriptive statistics. By reporting the numbers of adverse events (AE)/serious adverse events(SAE) within the trial to indicate the incidence rate of safety issues. We will also present the total number of patients suffering from AE/SAE and its percentage of all populations.

The incidence of adverse events will be tabulated, summarized by intensity and relationship to the investigational product (IP).

- All-Cause Mortality: A table of all anticipated and unanticipated deaths due to any cause, with number and frequency of such events in each arm/group of the clinical study.

- SAEs: A table of all anticipated and unanticipated SAEs, grouped by organ system, with number and frequency of such events in each arm/group of the clinical study. (See AEs definition below).

- Other (Not Including Serious) AEs: A table of anticipated and unanticipated events (not included in the SAEs table) within any subsets of the clinical study, grouped by organ system, with number and frequency of such events in each arm/group of the clinical study.

- ADEs: this category indicates the anticipated device-related adverse events which may be related to the digital stethoscope AS-101.

10.1 Adverse Events

The following list describes the common adverse events of procedural sedation (Non-ADE) and the device-related adverse event (ADE) to be recorded in this study.

Patient Problems	ADE	non-ADE	Definition	
Hypotension	-	AE	Decline for more than 30% from the Baseline over 1minutes	*
Bradycardia	-	AE	HR < 50 BPM	*
Tachycardia	-	AE	HR > 120 BPM	*
Hypoxemia	-	AE	Hypoxemia < 85 mmHg	O
Minor oxygen desaturation	-	-	SPO_2 < 94 Not reported	
Moderate oxygen desaturation	-	AE	SPO_2 < 90	*
Severe oxygen desaturation	-	AE	SPO_2 < 80	*
Hypertension	-	AE	Reach 30% over the baseline for more than 1mintutes	*
Apnea(not prolonged)	-	-	Apnea 15-30 second	*
Apnea(long)	-	AE	Apnea > 30 second	*
AW obstruction	-	AE		
Seizure	-	AE		*
Failed sedation	-	-		
Allergy	-	AE		
Cardiovasular collapse	-	SAE		

Cardiac arrest	-	SAE		
Myocardial infarction	-	SAE		
Aspiration Pneumonia	-	SAE		△
Prolonged Cognitive Dysfunction (> 6month)	-	SAE		△
Encephalopathy	-	SAE		△
Skin irritated allergic dermatitis	AE	-		
Skin irritation related allergenic shock	SAE			
			Red color is indicating “report to the IRB in a timely fashion”.	
REF: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5836671/			* Not include Transient Event （< 1mins)	
REF: https://pubmed.ncbi.nlm.nih.gov/27126387/			△ could lead to elongation hospital stay	
			O Hard to measurement	

Table 7. Adverse events

10.2 Deaths, Serious Adverse Events and other Significant Adverse Events

The serious adverse events of procedural sedation are also listed in the previous section and will be included in the safety analysis of this study.

10.3 Pregnancies

Pregnancy is one of the pre-anesthesia evaluation factors, in this study we’re not per-protocol excluding the pregnancies female.
The summary statistics will be produced in accordance with section 7.

10.4 Other Safety Measures

The summary statistics will be produced in accordance with section 9.” Vital signs might be appropriately included in this subsection. Many of the points made regarding laboratory tests in section 11 are relevant to vital signs.

11 Other Analyses

We will also record the vital signs and pre-determined events during the study as the following list.

Code	Event	Description
DIN	Injection of propofol	During the period of induction anesthesia.

DAD	Adding medicine	Medicine, dosing, Route of administration
OPS	Start procedure	Procedure type
OPE	End procedure	Procedure type
APN	Apnea	
JWT	Jaw thrust	
PAO	PAO	If more than 1 min, then record it every 1 min.
PTV	Patient actively vocalization	e.g., groan, speak, belch
FCV	Friction sound	Distinguishable friction sound on Airmod.
EHV	Exhaust sound	Distinguishable exhaust sound on Airmod
AAV	Non-patient abnormal sound	e.g., endoscope sound
CAV	Continuous abnormal sound	Wheeze, Stridor, Rhonchi
NCV	Non-continuous abnormal sound	Crackle, Saliva sound
EGV	Environmental massive noise	
PTM	Patient movement	Either actively or passively.
BDM	Bed start moving	
BDE	Bed stop moving	
BER	Device removement	
BEI	Device installation	
LBP	Low BP	Systolic BP < 30% baseline, individually decided by PI. [S/E]
HBP	High BP	Systolic BP > 30% baseline, individually decided by PI. [S/E]

HHR	Tachycardia	HR > 100, individually decided by PI. [S/E]
LHR	Bradycardia	HR < 60, individually decided by PI. [S/E]
IHR	Arrhythmia	Query PI about the type of Arrhythmia. [S/E]
ETF	EtCO2 failure	Unable to identify EtCO2 Signal. [S/E]
AMF	Airmod failure	Unable to read RR or Airmod crash. [S/E]
SPF	SpO2 failure	[S/E]
NSA	Nasal airway	[S/E]
OMO	On mask with O2	[S/E]
O2M	Modify O2	Record volume
AMP	Modify AS-101 AMP	Record AMP
ABB	Ambu bagging	
SAE	Serious adverse event	
DPS	Start dosing pump	Record dosing
DPM	Modify dosing pump	Record dosing and all action on dosing pump.
DPE	End dosing pump	
FIN	Discharge from PACU	PAR or PADSS score≥ 9

*[S/E]: The start time and end time of the event need to be recorded.
Table 8. Event recording list

12 Reporting Conventions

P-values ≥0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

13 References

Provide references for any citations in the main body of the SAP.

White IR and Thompson SG (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine,* **24,** 993-1007

Ramsay, M. A., Usman, M., Lagow, E., Mendoza, M., Untalan, E., & De Vol, E. (2013). The accuracy, precision and reliability of measuring ventilatory rate and detecting ventilatory pause by rainbow acoustic monitoring and capnometry. Anesthesia & Analgesia, 117(1), 69-75

14 Listing of Tables, Listings and Figures

Tables

Figures

Evaluating the Non-Inferiority of Airmod to Capnostream™35 on Respiratory Rate Monitoring and User Experience Clinical Evaluation

Statistical Analysis Plan

Comparison Table of Amendment

Clause Number	Current Version	New Version
9.1	The measurement AirRR is a time series list of respiratory rates value (rrValue) recorded by Airmod at 1Hz sampling rate from each subject.	The measurement AirRR is a time series list of respiratory rates value (**rrUI**) recorded by Airmod at 1Hz sampling rate from each subject.
9.1	The rrValue out of the claimed range (< 4 Brpm or > 35 Brpm) are excluded prior to efficacy analysis. The case number of the aforementioned record will be presented in the case report form.	The **rrUI** out of the claimed range (< 4 Brpm or > 35 Brpm) are excluded prior to efficacy analysis. The case number of the aforementioned record will be presented in the case report form.
9.2.2		**Also, we apply a sensitivity test for Airmod by comparing the reference method at each time point and calculating the absolute deviations in breaths per minute (bpm). As a measure of accuracy, absolute deviations were calculated as a percent of reference.** **Percent absolute deviation were averaged for each subject yielding overall measures of accuracy.** **To evaluate the accuracy of Airmod, the individual percent absolute deviation was compared for each monitoring time point by subtracting ManCRR from AirRR and then averaged within subject. A paired t test was used to evaluate significance across subjects.** **To measure the precision of Airmod for each subject, the standard deviation of percent deviations (in nonabsolute value) was calculated. The precision of Airmod was compared through a paired t test of the subject-specific standard deviations of Airmod's percent deviations from reference.**
13		**Ramsay, M. A., Usman, M., Lagow, E., Mendoza, M., Untalan, E., & De Vol, E. (2013). The**

		accuracy, precision and reliability of measuring ventilatory rate and detecting ventilatory pause by rainbow acoustic monitoring and capnometry. Anesthesia & Analgesia, 117(1), 69-75