STATISTICAL ANALYSIS PLAN

Protocol: CANARY01

Changes in Exhaled $^{13}CO_2/^{12}CO_2$ Breath Delta Value as an Early Indicator of Infection in ICU Patients

Protocol Version Date: 15 AUG 2016
SAP Version Date: 15 OCT 2016

Table of Contents

1. Abbreviations

AE	Adverse Event
BDV	Breath Delta Value
CDM	Clinical Data Management

2. Study Summary

The Isomark Canary™ is a device designed to measure the breath delta value (BDV) in exhaled breath samples obtained from a patient. The BDV is a ratio of the two primary naturally occurring isotopic forms of carbon dioxide, 13CO2/12CO2. Changes in the BDV can indicate the onset of infection due to the phenomenon known as the acute phase response, which occurs early in the infection process, even before a patient shows typical symptoms of an infection. Breath samples are collected in a disposable, single-use bag. The study is designed to collect data that can demonstrate the effectiveness of the BDV as an indicator of infection as compared to current methods, such as vital signs.

This is a prospective study using subjects who are not suspected of infection at the time of enrollment and comparing the Canary's BDV marker to standard infection monitoring methods.

3. Roles and Responsibilities

The Isomark team is involved in research (protocol) design and breath sample analysis.

Primary Study Contacts
Emily Breunig
Isomark – Clinical Project Manager
Office: (608) 561-1895 Mobile: (608) 438-5370
Email: emily.breunig@isomark.com
Contact with questions on protocol, enrollment, EDC, or to request additional supplies.

Stephanie Manthey
Isomark – Office Manager & EA
Office: (608) 561-1895 Mobile: (608) 852-9870
Email: stephanie.manthey@isomark.com
Contact with questions on the EDC or to request additional supplies.

Daniel Bütz, PhD
Isomark – Chief Scientific Officer
Office: (608) 561-1895 Mobile: (608) 358-5459
Email: daniel.butz@isomark.com
Contact with questions on the protocol.

4. Data Source

The purpose of this Statistical Analysis Plan (SAP) is to outline the processes in which the statistical considerations were made for this protocol. It will serve as documentation for the clinical data management procedures that will be followed for Isomark's CANARY01 trial.

Data will be collected from clinical sites on adult subjects that have been identified as critically ill who are at risk for developing an infection. Subjects will be enrolled based on the follwing criteria:

Inclusion criteria

1. Age 18 years or older
2. Critically ill patient admitted to the Intensive Care Unit (ICU)
3. High risk of infection, per Investigator discretion
4. Enrolled within 48 hours of ICU admittance (see Enrolled definition in section 6.1)
5. Expected duration of hospital stay at least 120 hours (five days) from time of study enrollment
6. Subject/LAR speaks a language of which the IRB has approved a consent form

Exclusion criteria

1. Confirmed infection at time of enrollment
2. Known use of non-prophylactic antibiotics, antimicrobials and/or antifungals for more than 24 hours prior to ICU admission
3. Currently active cancer, defined as receiving treatment or intend to receive treatment within hospital stay for cancer (including but not limited to: radiation, chemotherapy, systemic orals, etc)
4. If not intubated, unable to cooperate with providing a breath sample
5. Expected death within 24 hours of enrollment or lack of commitment to aggressive treatment by family/medical team (e.g., likely to withdraw life support measures within 24 hrs of screening)
6. Female who is pregnant or lactating (negative serum or urine pregnancy test results within 48 hours of enrollment or to be performed during screening)
7. Prisoner
8. Ethnic/Racial/Covariate category has completed enrollment
9. Known participation in an investigational and interventional research study within 30 days prior to enrollment, not approved in advance by sponsor (note: to be eligible, any interventional treatment must have ended at least 30 days ago)
10. Individuals who are directly affiliated with sponsor or study staff, or their immediate families. Immediate family is defined as spouse, domestic partner, parent, child, or sibling whether legally adopted or biological.
11. Any patient that is deemed unfit for study participation, per the Investigator's discretion

Data collected includes:

- Open enrollment of subjects, population defined by inclusion/exclusion criteria
- Subject breath sample taken at enrollment and every 4 hours for the duration of the subject's ICU stay
- SIRS/MEWS and vitals data
- Subject records monitored/collected daily
- Daily blood samples collected for the duration of the subject's ICU stay and stored for

subsequent analysis of CRP and PCT

5. Analysis Objectives

The overall objective of the analysis is to demonstrate that the BDV is a leading marker for infection diagnosis in critically ill subjects. Key questions to be answered are 1) Evaluate the BDV as a marker of infections in critically ill subjects at risk of progressing to severe sepsis and septic shock; 2) To measure variation of BDV with time in critically ill ICU patients progressing to severe sepsis and septic shock among those who agree to participate as research subjects; 3) To determine the magnitude of change of BDV in subjects who subsequently progress to severe sepsis and septic shock; 4) To define variation of BDV in critically ill ICU subjects who do not progress to severe sepsis and septic shock.

The BDV will be analyzed to determine the range of exhaled $^{13}CO_2/^{12}CO_2$ BDV in adult subjects. The $^{13}CO_2/^{12}CO_2$ BDV will be compared with current clinical standard scores, physiologic measurements and biomarkers (e.g. vital signs, WBC, CRP, PCT), and any interventions -- such as antibiotic administration -- will be noted for the purpose of assisting in understanding variation of BDV. For the purpose of defining a positive BDV for the presence of an infection, a positive BDV is defined as a change of greater than 1‰ within a 24-hour period above or below the baseline value. The 1‰ threshold for change in the BDV within 24 hours is based on the variability of approximately 0.5‰ seen in subjects with stable nutrition who do not develop infections. Thus, a change of 2 standard deviations away from the pre-infection mean will be examined as part of sensitivity among those with progression to severe sepsis and septic shock and positive predictive value for progression to severe sepsis and septic shock.

For critically ill patients admitted to an ICU, the mean overall nosocomial infection rate is 6 to 18% (29, 30, 40). The incidence rate of infection could be greater for all surgical ICU patients. The infection rate in the previous proof-of-concept study was 50% using the same population and inclusion/exclusion criteria; this assumption of equal proportions of prospectively enrolled infected and uninfected subjects will be used in the sample size calculations which follow. Data collected to date indicate that a sustained change in BDV of 1.0‰ or more correlates with the presence of infection when the subject is used as his-her control. The average intra subject standard deviation (SD) across time points in critically ill adult subjects previously studied was 0.95‰, regardless of the underlying medical conditions. The inter subject SD is expected to be 1.0‰ at most. Using the BDV measurements during the breath sample monitoring period (breath sample assessments every 4 hours), the expected overall SD of the mean BDV measurements (across time points) is less than 1.0‰.

BDV sensitivity and specificity will be evaluated relative to PCT sensitivity and specificity according to non-inferiority hypothesis tests where a greater than 10% BDV disadvantage is to be ruled for both sensitivity and specificity relative to a less than 10% BDV disadvantage. The BDV sensitivity and specificity estimates will then be used to determine the positive predictive value (PPV) and negative predictive value (NPV) once the prevalence is known; prevalence surrogates can be used to apply the observed sensitivity and specificity to various ICU and CCU settings.

<u>Objectives</u>
To determine the following:

1. Evaluate the BDV as a marker of infections in critically ill subjects at risk of progressing to severe sepsis and septic shock;
2. To measure variation of BDV with time in critically ill ICU patients at risk of progressing to severe sepsis and septic shock among those who agree to participate as research subjects;
3. To determine the magnitude of change of BDV in subjects who subsequently progress to severe sepsis and septic shock;
4. To define variation of BDV in critically ill ICU subjects who do not progress to severe sepsis and septic shock.
5. Compare sensitivity and specificity, and positive and negative predictive value of the PCT marker with the BDV.

Hypotheses

Co-Primary Endpoint 1

1. Null hypothesis: BDV sensitivity is more than 10% worse than PCT sensitivity for progression to severe sepsis and septic shock in critically ill adult ICU subjects
2. Alternative hypothesis: BDV sensitivity is less than 10% worse than PCT sensitivity for progression to severe sepsis and septic shock in critically ill adult ICU subjects

Co-Primary Endpoint 2

1. Null hypothesis: BDV specificity is more than 10% worse than PCT specificity for those not progressing to severe sepsis and septic shock in critically ill adult ICU subjects
2. Alternative hypothesis: BDV specificity is less than 10% worse than PCT specificity for those not progressing to severe sepsis and septic shock in critically ill adult ICU subjects.

Secondary Endpoint 1

1. Null hypothesis: the variability of BDV in critically ill adult ICU subjects does not change after clinical diagnosis of progression to severe sepsis and septic shock.
2. Alternative hypothesis: the variability of BDV in critically ill adult ICU subjects increases after clinical diagnosis of progression to severe sepsis and septic shock.

6. Analysis Sets/Populations/Subgroups

The first task is to demonstrate that there is a positive signal associated with the Canary test using the subject as their own control.

A per protocol analysis will be performed following EAC event reconciliation and subsequent database lock. The EAC will determine diagnostic truth by documenting severe sepsis or septic shock events as well as the onset and resolution dates.

Infection documentation will be reviewed by the EAC committee based on best clinical practices, using four categories:

- No suspicion of infection
- Inconclusive evidence of infection
- High suspicion of infection
- Overt infection

For sensitivity and specificity, non-inferiority testing will be performed to rule out a -10% disadvantage according to a one-sided 97.5% lower confidence limit for the BDV – PCT

difference; exact McNemar paired comparison tests will performed to test BDV non-inferiority relative to PCT. Both endpoints must achieve statistical significance in order to achieve study success.

PPV and NPV will be computed as observed and for a range of possible infection prevalence (10%, 20%, 30%, and 40%). Results will be analyzed in the same manner as sensitivity and specificity.

BDV variations will be assessed using a two-sample t-test to compare the averaged (across time points) BDV between subjects who are diagnosed with a severe sepsis and sepsis shock vs. no severe sepsis or sepsis shock. Also a linear mixed effects model with subject specific random effects will be used to compare BDV measurements between the two diagnostic groups. This model will be also used to estimate the intra- and inter-subject variability of the BDV measurements. A paired t-test will be used to compare the BDV measurements prior to and after a diagnosis of infection or sepsis.

The definition of a positive BDV will be based on a 72-hour window without advance knowledge of the diagnosis. In the absence of clinical details supporting PCT operating characteristics (sensitivity, specificity), two PCT thresholds (0.5 and 2.0) will be evaluated but the 0.5 threshold is deemed to be primary for BDV comparison.

Descriptive statistics will be generated for patient characteristics, demographics, and on-study events relating to subgroups (see Section 7.4). For continuous variables, descriptive statistics may include the number of subjects from whom data was used in the calculation, mean, standard deviation, median, minimum, and maximum; frequencies and percentages may be displayed for categorical data.

The predicate marker PCT sensitivity is estimated to be between 64-82% based on a recent review of the published literature (19); PCT sensitivity will be assumed to be 75% for sample size calculations. BDV sensitivity is expected to range between 80-90% based on the proof-of-concept study. The study is designed to have at least 80% power and one-sided 2.5% Type I error. A greater than 10% disadvantage for BDV vs PCT (null hypothesis) is to be tested against a less than 10% disadvantage for BDV vs PCT (alternative hypothesis). As shown in Table 1, this non-inferiority hypothesis can be tested with 97 documented infection events. The BDV advantage vs PCT resulting in a -10% one-sided 97.5% lower confidence bound (CB) is also displayed in Table 1; if the sensitivities are 76.5% for BDV vs 75% for PCT based on 97 pairs of infection events, then the one-sided 97.5% lower CB would be -10%, which would establish BDV non-inferiority relative to PCT.

Table 1: One-sided 97.5% Lower 97.5% Confidence Limit: Sensitivity Paired Differences (simulation)

	Power	**N-I**
Confidence level, 1-α (one-sided)	0.975	0.975
Expected difference, π_T - π_S, Δ_1	0.100	0.015
Proportion discordant, $\eta = \pi_{10} + \pi_{01}$	0.400	0.315
Proportion both yes, π_{11}	0.600	0.600
Lower limit for π_T - π_S, LL	-0.100	-0.100
Number of simulations	1000	1000
Random seed for simulations	123	123

Power (%)	89	51
n (number of pairs)	97	97

The predicate marker PCT specificity is estimated to be between 56-78% based on a recent review of the published literature (19); PCT specificity will be assumed to be 65% for sample size calculations. BDV specificity is expected to range between 70-80% based on the proof-of-concept study. The study is designed to have at least 80% power and one-sided 2.5% Type I error. A greater than 10% disadvantage for BDV vs PCT (null hypothesis) is to be tested against a less than 10% disadvantage for BDV vs PCT (alternative hypothesis). As shown in Table 2, this non-inferiority hypothesis can also be tested with 97 documented non-infection events. The BDV advantage vs PCT resulting in a -10% one-sided 97.5% lower confidence bound (CB) is also displayed in Table 2; if the specificities are 66.5% for BDV vs 65% for PCT based on 97 pairs of non-infection events, then the one-sided 97.5% lower CB would be -10%, which would establish BDV non-inferiority relative to PCT. A 10% non-inferiority delta with respect to both sensitivity and specificity ensures that the performance of BDV is comparable to PCT, considering the difference in sampling methods and the potentially greater convenience of breath sampling for BDV.

Table 2: One-sided 97.5% Lower Confidence Limit: Specificity Paired Differences simulation)

	Power	**N-I**
Confidence level, 1-α (one-sided)	0.975	0.975
Expected difference, $\pi_T - \pi_S$, Δ_1	0.100	0.015
Proportion discordant, $\eta = \pi_{10} + \pi_{01}$	0.400	0.315
Proportion both yes, π_{11}	0.100	0.100
Lower limit for $\pi_T - \pi_S$, LL	-0.100	-0.100
Number of simulations	1000	1000
Random seed for simulations	123	123
Power (%)	87	52
n (number of pairs)	97	97

We will also categorize infections per respiratory and non-respiratory infections within our data analysis to determine if there is an impact on the breath delta value.

We will also categorize samples taken from a ventilator versus free-breathing patient for analysis. To assess the impact of collecting samples from the ventilator versus free breathing subjects, we examined breath samples right before and right after ventilator weaning. Eleven of our subjects have been ventilated at the beginning of our study and then extubated at some point during the study. Out of the eleven subjects who had enough data for the analysis, none showed significant changes in BDV after extubation, meaning that there were no differences between ventilated and free breathing values.

All analyses will be conducted under GCP and FDA 21 CFR 812 standards. SAS (SAS Institute, Cary NC) and StatXact (CyTel, Cambridge MA) will be used to conduct all analyses. All statistical testing will be two-sided using a 0.05 p-value to reach statistical significance.

7. Endpoints and Covariates

We divided the covariates into two categories: the first is related to baseline covariates, such as underlying medical conditions, caloric intake, infection history, immune markers, being on antibiotics, and other concurrent therapies at baseline. Baseline medical conditions will be collected per the Charleston Comorbidity Index. The second category is related to on-study covariates (possibly outcome correlative) like procalcitonin, new antibiotics, new surgeries, new treatments, changes in diet, healing rates, and clinical signs and symptoms (e.g. fever, chills, WBC, breaths per minute, pulse, electrolytes).

While we recognize that some of the aforementioned covariates can affect sample carbon dioxide concentration (ventilator settings, or exogenous oxygen), any diluents will affect both 13CO2 and 12CO2 proportionally and will not disturb the 13CO2:12CO2 ratio. Similarly, covariates that compromise air exchange in the lung do not affect the ratio of 13CO2:12CO2 since there is no fractionating process during the exchange process. As long as the patient is producing CO2, there will be exhaled CO2 albeit at potentially lower concentrations in subjects with reduced air exchange capacity, but the relative ratio of 13CO2 to 12CO2 remains constant. These points being stated, we have modified our study design to collect the necessary covariate data to demonstrate that CO2 concentration does not affect the 13CO2:12CO2 ratio, by capturing the following information and the changes to these:

- **Ventilator settings** (Ventilator Mode, Oxygen Concentration, PEEP setting, Tidal Volume/Pressure, FiO2 PaO2 and Flow Rate)
- **Lung dysfunctions** (Baseline comorbidities and diagnosed lung dysfunctions)
- **Nutrition and metabolic status** (Type of Nutrition, Type of Line, and Composition)
- **Comorbidity** (COPD, diabetes, metabolic and lung dysfunction)

8. Handling of Missing Values and Other Data Conventions

The number of missing samples will be minimized, but some samples will be missed due to clinical care logistics. A complete case analysis will be the primary analysis. An imputation based analysis will be a secondary analysis. Specifically, the last observation carried forward (LOCF) and multiple imputation method based on the Markov Chain Monte Carlo technique will be utilized to perform the imputation based analyses. A sensitivity analysis will be conducted to compare the results of the complete case analysis with the results of the imputation based analysis.

Interim analysis plan

After completion of 25 subjects, an interim analysis will be done to ensure subject safety and to assess positive predictive value of the BDV for progression to severe sepsis and septic shock. If no subject progressing to sepsis or septic shock subjects had a positive BDV test, then additional enrollment will be stopped, and inclusion criteria will be reviewed. The study will not be stopped for superiority.

Subject safety will be assessed. If any significant risks to subject safety are identified, then additional enrollment will be stopped, and safety considerations will be assessed. During interim data analysis subgroups will be reviewed. Particular subgroups will be targeted for enrollment as necessary.

9. Statistical Methodology

Descriptive statistics will be generated. For continuous variables, descriptive statistics may include the number of subjects from whom data was used in the calculation, mean, standard deviation, median, minimum, and maximum; frequencies and percentages may be displayed for categorical data.

A 2-sample t-test will be used to compare the averaged (across time points) BDV between subjects who are diagnosed with an infection, without an infection, with sepsis and without sepsis. A linear mixed effects model with subject specific random effects will be used to compare BDV measurements between infected and non-infected and septic and non-septic subjects. This model will be also used to estimate the intra- and inter-subject variability of the BDV measurements. A paired t-test will be used to compare the BDV measurements prior to and after a diagnosis of infection or sepsis. Fisher's exact test and the Kappa statistic will be used to evaluate the relationship between a positive indication of infection by the BDV and a diagnosis of infection or sepsis in the subsequent 72 hours. Sensitivity, specificity, negative predictive value, positive predictive value and corresponding 95% confidence interval will be calculated.

10. Rationale for any Deviation from Pre-specified Analysis Plan

Not applicable at this time.

11. QC Plans

- Study personnel will complete sample collection and sample analysis records to ensure proper sample collection analysis and data reporting for each subject enrolled.
- When a new subject is enrolled in the study, the study team member will label and complete the sample collection records with the subject's study number.
- The study team member will notify Isomark personnel by email that a new subject is enrolled.
- At each subsequent breath collection time, a study team member will label the breath sampling equipment with the subject study number, site number, date and time when each sample was collected. The study team member will complete and initial the relative collection records within the Case Report Form (CRF). If the sample is obtained late, then the reason will be documented, e.g. subject in MRI. Each breath sample will be collected as close as possible and not exceeding one hour prior to or after the designated time for the sample collection.
- At each blood collection time, a study team member will label blood sampling equipment with collection date and time. Isomark will provide pre-printed labels with each subject's study number. The study team member will complete and initial the blood sample collection record, accounting for the time of collection, time of spinning, and time of freezing. The total time between collection and freezing cannot exceed four hours. If a sample is not collected on a given day, it will be considered missing data.

- A study team member will collect a breath sample every four hours and a blood sample daily for the study duration specified.
- Collected breath samples and the original copies of the corresponding sample collection records will be stored until analyzed on-site. Remaining sample will be shipped to Isomark.
- Blood samples and the original copies of the corresponding blood sample collection records will be placed in an Isomark-provided container and shipped to Isomark's contracted reference lab via an Isomark-approved carrier. The blood samples are processed by the reference lab and the results are sent to the sites, and will need to be entered into the EDC by the site staff. After the results have been successfully received, the lab will dispose of any remaining blood samples.
- Each time breath samples are received, labels will be checked for complete information and compare the samples to the EDC, noting anomalies. Any discrepancies will be cleared up between the site staff, contracted laboratory and Isomark personnel.
- At the beginning of each day on which samples are analyzed, the sample analysis technician will ensure the Canary is within calibration specifications by analyzing a standard sample of known isotopic composition and initial the sample analysis record.
- Breath samples are to be analyzed within 21 days of the time of collection. Isomark personnel will note the time and date of sample analysis on the sample analysis record and initial.
- Data will be recorded on the machine, and a verification code will populate the screen. The verification code will be recorded by the technician.
- Data that is extracted from the machine on regular intervals by Isomark designated personnel, will be recorded on an applicable log and cross checked to the technician's log. Discrepancies will be addressed and mitigated as soon as possible.

12. References

13. Appendices