Cover Page

NCT04269161

NICU Oxygen Control Study

Statistical Analysis Plan

October 21, 2024

Statistical Design and Power

A crossover design was chosen because premature infants should have fewer events as they grow older each day, and a crossover design will mitigate this potential confounding effect. Furthermore, because every subject receives the device and the standard of care (nurse only), every subject serves as their own perfectly matched control. All subject-specific effects cancel out, so there is no concern for covariate imbalance. Finally, selection bias is minimized because all subjects receive all "treatments". Nonetheless, the concealed allocation will be used with blocked randomization at each site, using sealed envelopes. The subjects will be randomized to group A or B in sets of 8 (i.e. in each group of 8 envelopes, 4 will be group A and 4 will be group B). During the entire study process the infants will receive normal NICU care and the parameters the SpO_2 range will be set by the physician caring for the infant. There are also built in manual overrides for the device, which allow the NICU staff to make changes while the subject is on the device phase of the study. The device will be able to record these changes and the staff will record their manual interventions in the study diary.

We have planned our sample size using a non-inferiority test for a 24 period, 2 treatment cross-over design, based on our primary endpoint, t_Δ. For a given patient, define t_Δ = (mean elapsed time needed for device to re-establish SpO_2 after alarm) - (mean elapsed time needed for nurse to re-establish SpO_2 after alarm). The margin of non- inferiority will be chosen as $t_\Delta > -10$ sec, so that a device which is no worse than 10 sec, on average than a nurse will be considered non-inferior. Assuming the standard deviation of $t_\Delta = 12$ and the true mean difference is zero under the alternate hypothesis, a sample size of 48 achieves 88% with $\alpha = 0.05$. If there is 16% patient drop- out before crossover, so that the final n=40, the power drops to 82%. In all analysis, a (paired) t-test will be used. Our secondary endpoint will be analyzed in a similar manner. The secondary endpoint is the proportion of time SpO_2 is within the prescribed range, using an area-under-the-curve approach (with a discrete state) to account for varying time-on-test.

We will plan for one interim analysis to determine if the trial should be stopped early due to futility (strong evidence of inferiority, where a confidence interval for t_Δ lies entirely to the left of -10 and doesn't intersect -10) or for efficacy (strong evidence of superiority with margin > +20 sec). This will be carried out when n=32 (16 subjects per site) is attained and stopping decisions will be based on O'Brien-Fleming stopping principles. The interim analysis will be carried out by an independent statistician on the University of Missouri's Data Safety and Monitoring Committee, which is also available to monitor the study for adverse events if requested by the IRB. In the event that the patient drop-out is greater than 16% before crossover, then a more complicated estimation procedure will be employed using mixed effects models; otherwise, complete cases will be used.

Assuming that the probability of enrollment for a given NICU patient is 10% (conservative est.), and because there are ≥6 new NICU patients/week at each hospital, we expect to enroll 48 patients and collect data in 43 weeks (11 months).