Statistical Analysis Plan

Implementation of **N**udges to **P**romote **U**tilization of low **T**idal volume ventilation (INPUT) Study

Sponsor	**National Heart, Lung, and Blood Institute**
IRB Protocol Number	**833400**
Principal Investigator(s)	**Meeta Kerlin, MD, MSCE**
Data Analysts	**Wei Wang, PhD** **Yingying Lu, PhD**
Clinical Trials.gov Number	**NCT04663802**
Version Date:	**March 1, 2024**

I. BACKGROUND AND RATIONALE 3
II. OBJECTIVES AND HYPOTHESES 3
III. DESIGN AND RANDOMIZATION 3
IV. POPULATION 4
V. PRIMARY OUTCOME AND ANALYTIC METHOD 4
A. PRIMARY OUTCOME 4
B. PRIMARY ANALYTIC SAMPLE 4
C. PRIMARY ANALYSIS 5
VI. SECONDARY ANALYTIC SAMPLE 6
VII. ANALYSES OF PATIENT-LEVEL EFFECT MODIFIERS 7
VIII. SUBGROUP ANALYSIS 7
IX. SENSITIVITY ANALYSIS 7
X. SECONDARY OUTCOMES 8
A. ANALYSIS OF SECONDARY OUTCOMES 9
XI. APPROACH TO MISSING DATA 10
XII. SAMPLE SIZE AND STATISTICAL POWER CALCULATIONS 10
A. SAMPLE SIZE 10
B. STATISTICAL POWER CALCULATIONS 11
XIII. DATA SAFETY & MONITORING BOARD (DSMB) 11
XIV. SAFETY DATA AND INTERIM ANALYSIS 12
A. SAFETY MONITORING ENDPOINTS 12
B. TIMING OF ANALYSIS AND ADJUSTMENT OF SIGNIFICANCE LEVEL 12
C. TIMING OF FINAL ANALYSIS 12
D. PERSON PERFORMING ANALYSIS 12
XV. REFERENCES 14
APPENDIX A 15

I. Background and rationale

Of the million Americans who undergo invasive mechanical ventilation (MV) annually within intensive care units (ICUs), 30-35% die in the hospital, and survivors commonly have long-term cognitive, emotional and physical impairments. Although potentially life-saving, MV can also injure the lungs. Strategies of "lung-protective ventilation" (LPV) aim to minimize harm by delivering low tidal volumes (the amount of air per breath delivered by the ventilator) and limiting the artificial pressures in the lungs. LPV was first proven effective in patients with acute respiratory distress syndrome (ARDS), a severe form of respiratory failure associated with pneumonia, sepsis, and other common illnesses. A multicenter randomized trial published in 2000 demonstrated an absolute mortality reduction of 10%, shortened duration of MV, and reduced organ failure rates. These findings have since been replicated and incorporated into international guidelines. However, fifteen years later, more than one-third of ARDS patients overall, and up to 81% in some ICUs, do not receive LPV.

Multiple studies across diverse ICUs have consistently identified two key barriers to LPV utilization: limited knowledge about LPV, and tendencies to prescribe LPV only when ARDS is definitively diagnosed, in part due to theoretical concerns about potential harms of LPV. However, definitive diagnosis of ARDS is challenging. And recent evidence suggests that (1) concern for harm among patients without ARDS is likely unwarranted, and (2) LPV may in fact reduce lung injury and mortality even among patients without ARDS.

II. Objectives

The trial's overall objective is to compare simple, low-cost, scalable strategies grounded in behavioral economic theory to (a) increase utilization of evidence-based LPV for MV patients, (b) gauge the process-oriented and clinical benefits of targeted vs. non-targeted approaches to changing behavior, (c) explore how clinician and environmental contextual factors affect LPV utilization, and (d) determine how patient factors such as type and severity of disease affect both LPV utilization and its associated outcomes.

Overall, we hypothesize that strategies that encourage clinicians to use LPV among all MV patients will mitigate the aforementioned barriers of knowledge gaps and diagnostic uncertainty, thereby increasing LPV utilization across diverse settings among patients with and without ARDS.

III. Design and randomization

- Stepped-wedge, cluster-randomized, pragmatic clinical trial of three EHR-based implementation strategies designed to increase utilization of LPV among MV patients. (see **Appendix A**).
    - 12 ICUs across five UPHS hospitals, which all currently have access to an EHR-based algorithm that identifies patients with ARDS and prompt physicians to employ LPV via a clinical dashboard and through ICU telemedicine support, will sequentially add two of three EHR-based implementation strategies to further promote LPV utilization among all MV patients.
    - Study ICUs will be randomly assigned to first receive one of two physician-directed strategies: either a **default order set (Strategy A)** or **physician-targeted accountable justification strategy (Strategy B)**. ICUs will be assigned to one of six wedges, thereby determining the date on which they adopt their assigned EHR-based strategy (**Figure 3**). The first wedge will begin in the fourth month of the trial phase. Six months after adoption, all ICUs will add on an **RT-targeted accountable justification strategy (Strategy C)**. The trial monitoring period will last 27 months. This design enables comparisons of outcomes before and after implementation *within* ICUs, as well as at a given point in time *among* ICUs which will have been randomly assigned to different strategies.
    - All ICUs will contribute a minimum of 3 months of baseline data prior to intervention and utilize two strategies in combination for a minimum of 3 months.

- After the 27-month trial monitoring period, all study ICUs will contribute 6 months of post-intervention data for analysis of sustainment, during which time the research team will have no further contact with the study ICUs. Therefore, the study intervention period will be 27 months, and the entire study duration will be 33 months.
- 12 ICUs across five UPHS hospitals will be randomized into six 2-ICU clusters by computerized random-number generation using R random number generator function.
- To form clusters with balanced patient volume and primary outcome baseline value, ICUs will be matched in pairs based on: 1) ICU patient volume (ICUs will be categorized as large or small based on whether their annual patient volume is above or below the median) and 2) the baseline value of the primary outcome (ICUs will be categorized as high or low rates of adherence to LPV based on whether their median rate is above or below the median). The six ICU clusters will then be randomly assigned to receive either Strategy A or B and then randomly assigned to represent wedges 1 through 6, with each wedge transitioning to the intervention phase in sequential fashion.
- We selected the timing of each wedge transition such that each wedge (2-ICU cluster) contributes a minimum of 12 weeks of data under the control condition prior to intervention implementation.

IV. Population

The trial will have two subject populations.

- The first group includes mechanically ventilated patients admitted to study ICUs during the pragmatic trial of interventions that will be applied at the level of the ICU.
- The second population includes mechanically ventilated patients admitted to study ICUs in the 6 months after intervention phase has been completed. During this observational phase, there will be no additional experimental interventions or contact with the study team.

V. Primary outcome and analytic method

a. Primary outcome

The primary outcome variables will be (1) **fidelity** to LPV, measured by the percentage of time in the first 72 hours of mechanical ventilation that a patient is exposed to tidal volume $\leq$6.5 cc/kg ideal body weight and (2) **sustainability** of the effects of the implementation strategies, measured as the percentage of time in the first 72 hours of MV that a patient is exposed to tidal volume $\leq$6.5 cc/kg ideal body weight among MV patients each month during 6 months after the study monitoring period ends.

We will consider episodes separated by less than or equal to 2 hours as a single episode. We will count the period of time that a patient is extubated as if the patient is on minimal settings, and will therefore subtract that period of time from the denominator of the primary outcome variable (if it falls in the first 72 hours).

b. Primary hypotheses

Hypothesis 1: The default order panel will increase utilization of LTVV compared to usual care.

Hypothesis 2: The accountable justification order panel will increase utilization of LTVV compared to usual care.

c. Primary analytic sample

For the purpose of the trial, the **primary analytic sample** will be limited to episodes of MV among patients who meet the following criteria:

1. Aged 18 and over; **AND**
2. Admission to 1 of the 12 participating ICUs; **AND**
3. Undergoing mechanical ventilation

Any episodes of MV that meet the eligibility criteria above will be **excluded** from the primary analysis if they meet any of the following criteria:

1. The episode of MV lasts less than 12 hours, because we believe that the evidence-based practice may not apply to these patients nor alter their outcomes.
2. The patient is on minimal settings for the entirety of MV, defined as a spontaneous mode (e.g., pressure support ventilation) with pressure support $\leq$10 cmH2O, **AND** PEEP $\leq$8 cmH20, **AND** FiO2 $\leq$50%, because the clinical significance of spontaneous tidal volumes is unknown and low tidal volumes may not be beneficial or desirable.
3. Goals of care are documented as comfort measures only (as identified by presence of an order for "Comfort Care" in the EHR) during the first 72 hours during episode of MV, because mechanical ventilation is managed differently during care focused exclusively on comfort and low tidal volume ventilation may not be appropriate, nor would it likely influence clinical outcomes.
4. There is no height documented in the EHR at the time of initiation of MV, because we will be unable to estimate ideal body weight, a necessary parameter to calculate the primary outcome, and because they will not receive the interventions.
5. The height documented is less than 4 feet, because the formula for ideal body weight does not hold true below this height.

d. Primary analysis

The primary analytic approach will use a **modified intention-to-treat (mITT)** approach. We choose this approach because we have observed in preliminary studies in the study ICUs that some patients will start MV and receive MV orders prior to ICU admission, and because of the constraints of the EHR to deploy the strategies, such episodes will not be eligible to be exposed to the intervention. Because ICUs differ in their percentages of patients who will and will not be eligible for exposure to the strategies, and these patient groups may be systematically different, in the primary analysis, we will limit the population to patients who receive a new MV order within 72 hours of enrollment (the time of the start of MV in the ICU).

The unit of analysis for primary study outcome (and all implementation outcomes) will be the episode of mechanical ventilation.

- Patients may have multiple episodes of MV if the patient is ventilated multiple times. We will include only the first eligible episode during a hospital admission for the primary analysis. We will calculate the proportion of patients who receive multiple episodes or multiple admissions. We expect the proportion will be very small, based on prior data from study ICUs.

The primary comparisons of interest will be:

- strategy A vs. strategy B
- strategy A vs. control
- strategy B vs control
- strategy A vs. strategy A combined with strategy C
- strategy B vs. strategy B combined with strategy C

Secondary comparisons of interest are control versus A + C, control versus B + C, and control versus C alone among patients who are not exposed to the order panel strategies (i.e., in the ITT but not MITT population).

We will use a linear mixed-effects model for primary analysis. The structure of the model is as follows:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + X_{ij}\theta + e_{ijk}$$

where Y_{ijk} denotes the percentage from individual *k* at time *j* from cluster *i;* α_i is a random effect for cluster *i*; X_{ij} is the treatment indicator for cluster *i* at time *j,* θ is the treatment effect and the residual is *iid* with $\mathrm{e_{ijk}} \sim N(0, \sigma_e^2)$. Since the intervention occurs over time, the proportion of clusters exposed to the intervention gradually increases. We will include step as a fixed effect (β_j) in the model to adjust for the potential confounding factor from calendar time.

In addition, to account for the matched pair randomization, which was done to balance patient volume and estimated baseline adherence to low tidal volume ventilation per cluster, we will include variables for the total number of observations within a cluster per step and estimated baseline adherence rates per cluster.

All analyses will be conducted with adjustment for pre-specified cluster-level and patient-level covariates that exist prior to the patient meeting eligibility criteria (**Table 1; see Appendix B for additional details**). We will include the adjustment terms $\varphi_1 \mathrm{Z_{ijk}}$ and $\varphi_2 \mathrm{W_{ij}}$, where ***Z*** and ***W*** represent vectors of patient and cluster characteristics that could be predictors of outcomes, to adjust for any imbalance from clusters' and patients' characteristics. Note that the index *j* in these matrices allows us to include time-varying covariates. To account for the possibility of a delay in an intervention exerting its effects, we will change $X_{ij}\theta$ to $X_{ijl}\theta_l$ to account for the number of months since the intervention was introduced in cluster *i*. We intend to include these additional terms (with pre-specified patient and ICU characteristics) in our primary models to augment precision in estimating the treatment effects.

Table 1. Factors for adjustment
Age
Sex
COVID-19 status (categorical)
LAPS2 severity of illness measured at the time of hospital admission
LAPS2 severity of illness measured at the time of enrollment
ICU admission source
Hospitalization duration prior to ICU admission
Code status (categorical)
Location of start of MV episode
Height

If the fitted model violates the 'normality of residuals' assumption, we will use a suitable inference approach such as permutation test or bootstrap confidence intervals. In addition, we will explore alternative modeling approaches to account for non-normality if needed, e.g. two part mixed effects models.

All analyses will be conducted using R (Vienna, Austria).

VI. Secondary analytic sample

We will repeat the primary analyses using two secondary analytic samples. First, we will include the true ITT population; that is, the entire population of episodes of MV in any study ICU (after exclusions detailed above), regardless of whether or not there was exposure to an implementation strategy during the intervention period. This analysis will provide insight into how ICU-targeted strategies may influence the strategies' effectiveness. We hypothesize that the effect of the strategies to improve administration of low tidal volumes will be attenuated in the true ITT population, compared to the modified ITT population. Secon, we will perform a per protocol analysis, restricting the study population to only those episodes where patients were actually exposed to the order panel nudges in the intervention groups and patients who received new orders for mechanical ventilation after the time of ICU transfer in the control group (to identify equivalent patients in the usual care population).

VII. Analyses of patient-level effect modifiers

Patient-level effect modification will be explored by repeating the analyses of the primary outcomes stratified by four candidate effect modifiers:

(1) presence or absence of ARDS
(2) degree of hypoxemia (as measured by the P:F ratio, calculated as the PaO_2 divided by the fraction of inspired oxygen (FiO_2)).
(3) Race/ethnicity
(4) Sex

If differences appear in stratified analyses, we will formally evaluate for effect modification by testing the significance of coefficients for statistical interaction terms between the potential effect modifier and the study groups on the primary outcome of LPV utilization, and on a subset of effectiveness outcomes: in-hospital mortality and hospital length of stay.

VIII. Subgroup analyses

In addition to the analyses above of patient-level effect modifiers, we will perform the following pre-specified subgroup analyses:

1. Patients without COVID-19, because we have observed that adherence to low tidal volume ventilation is high among patient with COVID-19 and there will likely be little or no effect from the study interventions.
2. Patients in the ITT group who are excluded from the MITT group. These patients would be eligible for exposure only to the RT-targeted intervention, as by definition they would not have been exposed or eligible for exposure to Strategy A or B if they did not have a new order for MV placed after ICU admission. In this subgroup, the exposure variable will be exposure to Strategy C (based on the time of admission to the study ICU relative to the ICU's intervention status with regards to Strategy C).

IX. Sensitivity analyses

a. **Time effect might differ across clusters:** To account for the possibility that the time effect might not be the same for all clusters, we will add a random interaction term for the cluster-time combination φ_{ij} .
b. **Treatment effect might differ across clusters:** To evaluate the possibility of treatment-effect heterogeneity, we will add a random interaction term $X_{ij}\tau_i$ which allows different treatment effects for different clusters.
c. **Exclusion of subsequent MV episodes in patients with multiple MV episodes may lead to bias estimates of effect:** To address potential bias of excluding subsequent episodes of MV among patients who undergo multiple episodes, we will repeat the primary analysis including all eligible episodes of MV.

d. **ICU may be a potential confounder**. Because the study ICU may itself be a potential confounder, we will repeat analyses modeling the study ICU as a fixed effect.

In order to preserve an overall family wise error rate of 0.05, we will pre-specify a type I error rate threshold for the comparison of A vs. B at 0.025, and will estimate the Holm thresholds for the remaining comparisons by ordering unadjusted p-values (from the smallest to largest), and then calculating it as 0.025 divided by n remaining contrasts.

X. Secondary Outcomes and Potential Adverse Effects

We will analyze several secondary measures of implementation, effectiveness, and potential adverse effects, as listed in **Table 2**.

Table 2 Secondary outcome measures

Implementation Outcomes	***Variable coding***	***Definition***
• Total duration of exposure to tidal volume >8 cc/kg and >10 cc/kg IBW	Continuous	In hours; exclude time on minimal settings; report for first 72h after enrollment and for entire duration of MV episode within a study ICU
• Initial tidal volume administered	Continuous	In cc/kg IBW; exclude time on minimal settings; report the first documented tidal volume at or after the time of enrollment
• Plateau pressure (Pplat) >30cm H20 at time 0 and time 24h	Binary	Time 0 is the earliest measure between 0 and 6h; time 24 is the measure between 18h and 30h that is closest to 24h and earliest if there are two measures of the exact same distance from 24h (e.g., one before and one after)
Effectiveness Outcomes		
• Hospital mortality	Binary	Vital status at hospital discharge
• Hospital discharge disposition	Categorical	To be collapsed into 4-6 categories
• Duration of MV	Continuous	From time of enrollment to end of MV in study ICUs, up to 30 days
• ICU length of stay	Continuous	From time of enrollment to end of admission to enrolling ICU, up to 30 days
• Hospital length of stay	Continuous	From time of enrollment to time of hospital discharge, up to 30 days
Potential Adverse Effects		
• Early deep sedation	Continuous	Proportion of time during the first 72 hours of mechanical ventilation that patients were alive, in the ICU, and with Richmond Agitation-Sedation Scale (RASS) of -3 to -5

• Average sedation intensity within the first 72 hours	Continuous	Average RASS value, weighted by duration of time at that value
• Deep sedation for the entirety of the first 72 hours of mechanical ventilation	Binary	Whether the patient was kept at RASS -3 to -5 for the entirety of the first 72 hours of mechanical ventilation

We will define duration of MV per episode as:

[Time 1: time of discontinuation of MV, up to 30 days after Time 0] – [Time 0: time of first eligibility]

We will define ICU length of stay as:

[Time 1: time of discharge from study ICU, up to 30 days after Time 0] – [Time 0: time of enrollment]

We will define hospital length of stay as:

[Time 1: time of hospital discharge, up to 30 days after Time 0] – [Time 0: time of enrollment]

Measures of excessive sedation as a potential adverse consequence of LPV will be defined as above in **Table 2.**

a. Analysis of secondary outcomes

Secondary implementation outcomes will be analyzed at the level of the episode and will be modeled as per the primary analysis, using generalized linear mixed effects models, with appropriate links for binary and categorical dependent variables.

For example, for a binary outcome, the model has the form

$$logit(Y_{ijk}) = \mu + \alpha_i + \beta_j + X_{ij}\theta$$

Model parameters are similar to those in the primary outcome analysis.

Effectiveness outcomes and potential adverse effects will be analyzed at the level of the episode of MV. Binary and categorical effectiveness outcomes (mortality and discharge disposition) will be modeled per the primary analysis.

For MV duration, ICU length of stay, and hospital length of stay, because of a high mortality rate (~30%), we will model these outcomes as a composite with death using a “free days” approach. In other words, for duration of MV, for example, we will use ventilator-free days within 30 days, which will be calculated as the number of days a patient is alive and not on invasive mechanical ventilation for 30 days from the start of enrollment. We will use similar calculations for ICU-free days (alive and not admitted to the study ICU) and hospital-free days (alive and not admitted to the study hospital). For patients who are discharged alive and off the ventilator prior to 30 days, we will assume they are alive and off the ventilator for the remainder of 30 days. For patient who are discharged on the ventilator prior to 30 days, we will assume they remain on the ventilator for the remainder of 30 days. For patients who are discharged to hospice (home or facility), we will assume they have died upon the day of discharge.

We will explore the population where patients have multiple episodes of MV, as patients with multiple admissions may have higher risk of mortality. We will also examine the patient outliers in terms of MV

duration, ICU length of stay and hospital length of stay. Depending on the proportion, we may include multiple episodes into the study by creating a variable about the number of episodes.

XI. Approach to missing data

Based on preliminary analyses of missingness within a retrospective cohort of MV patients admitted to 11 of 12 study ICUs, we propose the following approaches to missing data, based on specific data elements:

- Missing or erroneous value for ideal body weight: Less than 5% of patients in a retrospective patient cohort including 11 of 12 study ICUs have either missing or erroneous data for patient height, which results in a missing or erroneous value of ideal body weight. Because the EHR-based implementation strategies will not function for patients with missing data for height (i.e., there will be no pre-populated value for tidal volume, or threshold to trigger the accountable justification prompts), we will exclude these patients from the primary analysis and will report the percentage of missingness in the results. We will also examine if there is any pattern of erroneous data.
- Missing data for ventilator settings: Ventilator flowsheet data are entered at variable intervals, depending on clinical circumstances (e.g., if a patient's settings are changed) and routine practice (e.g., the ICU documentation standards). Therefore, there are no regular intervals during which ventilator flowsheet data will be entered. At any given time point, some or all relevant ventilator data may or may not be entered by a respiratory therapist or nurse. If ventilator data settings are partially missing for a given time (e.g., one setting or parameter is documented but others are not at a certain time), we will use single imputation, carrying forward data from the prior time of documentation for any relevant parameters that are missing. We choose this strategy because we suspect that if a setting or parameter is not documented at a subsequent time, then that setting or parameter is unlikely to have changed from prior (i.e., a change in a setting or parameter would typically prompt documentation).
- Missing or erroneous data for potential confounders, including severity of illness measures: In our previous experience working with data from the retrospective cohort, we anticipate minimal missingness for variables that will be included in the severity of illness measure (LAPS2). In the rare event that we do have missingness for a parameter, we will use single imputation of a normal value for that parameter, as is the common approach in calculation of severity of illness meaures. This strategy would bias the results towards the null. We will report percentage of missingness and discuss the risk of bias in the limitations.

XII. Sample size and statistical power calculations

a. Sample size

The first group includes mechanically ventilated patients admitted to study ICUs during a pragmatic trial of interventions that will be applied at the level of the ICU. Patients will not be approached directly for inclusion in this study. The interventions will be embedded within the electronic health record that is in use at each hospital. The sample size will be determined by the number of eligible patients that are admitted to study ICUs during the study period, which we estimate will be 27 months during the trial monitoring period. An estimated 6000 patients undergo MV annually within 13 ICUs of the 6 hospitals of UPHS. By extrapolating these data to a 27-month enrollment period, excluding one ICU for strategy development and pilot- testing, and excluding a fraction of patients who will meet exclusion criteria specified above, we estimated before the start of the study that approximately 8,000 MV patients would be eligible for inclusion in the trial across the remaining 12 ICUs during the study's intervention period.

At the time of the first interim analysis , based on the rate of accrual during the first 9 months of the trial and with application of more precise exclusion criteria, we found that the number of eligible patients admitted to study ICUs was lower than projected. Therefore, we revised our progjects to estimated samples sizes of 5100 for the MITT (primary) population and 6900 for the ITT population.

The second group includes mechanically ventilated patients admitted to study ICUs in the 12 months after intervention phase has been completed. During this observational phase, there will be no experimental interventions. Study ICUs will be free to design and implement processes of care entirely at their discretion. Patients will be included in the analysis if they meet eligibility criteria, which will be identical to those of the first interventional phase of the study. By extrapolating estimated enrollment data to a 12-month enrollment period, we estimate approximately 3,000 MV patients.

b. Statistical power calculations

We estimated the power of pairwise comparisons of any of the intervention groups (Strategy A or B alone or in combination with C – i.e., four separate intervention groups) with the control group. We initially based our power calculations on the following conservative assumptions: sample size of 8,000 episodes across 12 ICUs; an intracluster correlation within ICUs of 0.1; and a baseline mean value of the primary outcome of 45% with a standard deviation of 45% (based on estimates from a retrospective cohort of patients admitted to ICUs in study hospitals during a 6-month period in 2020, including patients with COVID-19). With these assumptions, we estimate that we will have >95% power to detect an increase of 25% in the mean value of the primary outcome (from 45% to 70%), which would approximate the utilization rates of the study ICU with the highest adherence to tidal volume < 6.0ml/kg ideal body weight. This sample size also has >90% power to detect a difference of 20% and nearly 80% power to detect a difference as low as 15% in any pairwise comparisons. Finally, this sample size has 80% power to detect a reduction in in-hospital mortality from an estimated 25% to 16%.

With the new sample size estimates as detailed above, and using the same assumptions for revised power calculations, we estimated that we have at leat 97% power to detect an increase of 25% in the mean value of the primary outcome (from 45% to 70%) and greater than 80% power to detect an increase as low as 17% in the mean value of the primary outcome within the MITT population. In addition, the new estimated MITT sample size has approximately 80% power to detect an 11% reduction in hospital mortality from an estimated 25% to 14%.

XIII. Data Safety & Monitoring Board (DSMB)

An important safeguard to protect research participants is the development of a plan for ongoing data and safety monitoring to anticipate, and protect against, any human subjects research concerns that may arise. A convened DSMB comprised noted experts in critical care, palliative care, and biostatistics to guard against the possibility of any unforeseen risks arising during the study. Members of the DSMB will not be involved in the conduct of the trial. Once convened, the DSMB will perform several duties. First, they will review and approve the research protocol and plans for data and safety monitoring prior to the study's implementation. Second, they will evaluate the progress of the trial. This will include assessment of data quality, participant accrual and retention, participant risk versus benefit, and study outcomes. Third, they will make recommendations to ensure that any identified issues are appropriately addressed. The DSMB will meet after 11 months and 20 months (one-third and two-thirds of the way through the trial enrollment.) We will charge DSMB members with using their judgment in simultaneously considering many data points in making decisions about trial design modifications and trial continuation or termination. The PI (Dr. Kerlin), the project manager, and the data analyst (Dr. Wang), will participate in all DSMB meetings as non-voting members. The PI, assisted by the project manager, will be responsible for maintaining communication between the DSMB, the IRB, and study sites.

The trial's DSMB members to monitor the scientific conduct of the study are as follows:

- **Todd Rice**, **MD, MSc** Associate Professor of Medicine in the Division of Allergy, Pulmonary and Critical Care Medicine at Vanderbilt University (DSMB Chair)
- **Daniella Meeker, PhD** Assistant Professor of Preventive Medicine and Director of Clinical Research Informatics at the Keck School of Medicine University of Southern California
- **Kristin Riekert, PhD** Professor of Medicine in the Division of Pulmonary and Critical Care

Medicine, Department of Medicine at Johns Hopkins University and Director of the Johns Hopkins Adherence Research Center (JHARC) and the Director of the Cystic Fibrosis Adherence Program
- **Jing Cheng, MD, MS, PhD** Professor of Biostatistics within the University of California San Francisco Division of Oral Epidemiology & Dental Public Health and Division of Epidemiology and Biostatistics

XIV. Safety data and interim analysis

a. Safety monitoring endpoints

The trial has no plans to stop the trial early for evidence of effectiveness of the implementation strategies because doing so would reduce our power for secondary analyses and analyses of effect modification. We will propose to stop the trial for early evidence of harm based on in-hospital mortality.

b. Timing of analysis and adjustment of significance level

Two interim analyses will be conducted after approximately 9 months and 18 months after trial launch (one-third and two-thirds of the trial duration, respectively) of two study outcomes: (1) fidelity to LPV (primary outcome, as specified above), and (2) in-hospital mortality. All interim analyses will be performed at the level of an episode of mechanical ventilation. For the open portion of the DSMB meeting, we will summarize data for the entire study population in aggregate, including patient characteristics, key patient outcomes, and the primary study outcome (percentage of time during the first 72h of MV in the ICU that the patient has tidal volume less than or equal to 6.5cc/kg IBW). We will also report adverse event rates among in aggregate.

To evaluate mortality by group, we will first report unadjusted mortality rates by group, and then proceed in a stepwise fashion with up to three risk-adjusted models:

- Model 1: We will use a generalized linear mixed effects model to estimate unadjusted mortality accounting for design issues; that is, including a random effect for ICU, a fixed effect for study group assignment, and a fixed effect for step (to account for potential confounding from calendar time).
- Model 2: We will use a generalized linear mixed effects model to preliminarily test whether study group assignment is associated with mortality after risk adjustment. As in the above model, we will include a random effect for ICU, a fixed effect for study group assignment, and a fixed effect for step (to account for potential confounding from calendar time). Given variability in case mix, we will also include the following covariates for risk adjustment: age, sex, race, source of admission to the hospital, Elixsauser score and LAPS2 score on the day of study enrollment.
- Model 3: If in the risk-adjusted mixed effects model we find an estimated risk difference for mortality of greater than or equal to 5% at a significance level of $p<0.2$, where either intervention group has higher mortality compared to the control group, then we will proceed with further analyses. Because interim analyses of accumulating data in a clinical trial can inflate type I error, we will use a flexible spending function. Specifically, we will calculate a Lan-DeMets spending function using O'Brien-Fleming boundaries to preserve the overall significance level of 0.05. The interim tests will use one-sided significance levels of 0.0007 and 0.0161. Final analyses will be conducted at a significance level of 0.0451, preserving a trial-wide Type I error rate of 0.05.

c. Timing of final analysis

Final analysis of the trial monitoring period will occur following completion of the trial in May 2023.

d. Person performing analysis

Statistical analysis will be performed by the trial's data analyst who will be blinded to trial arm.

All other members of the investigative team except the data manager and project manager will remain blinded to treatment assignments and facilities for the analysis. The trial's data manager and project manager will be the only people who knows the treatment assignments during the trial's conduct. Neither the data manager or project manager will be involved in developing or modifying the statistical analysis plan or in the actual analyses.

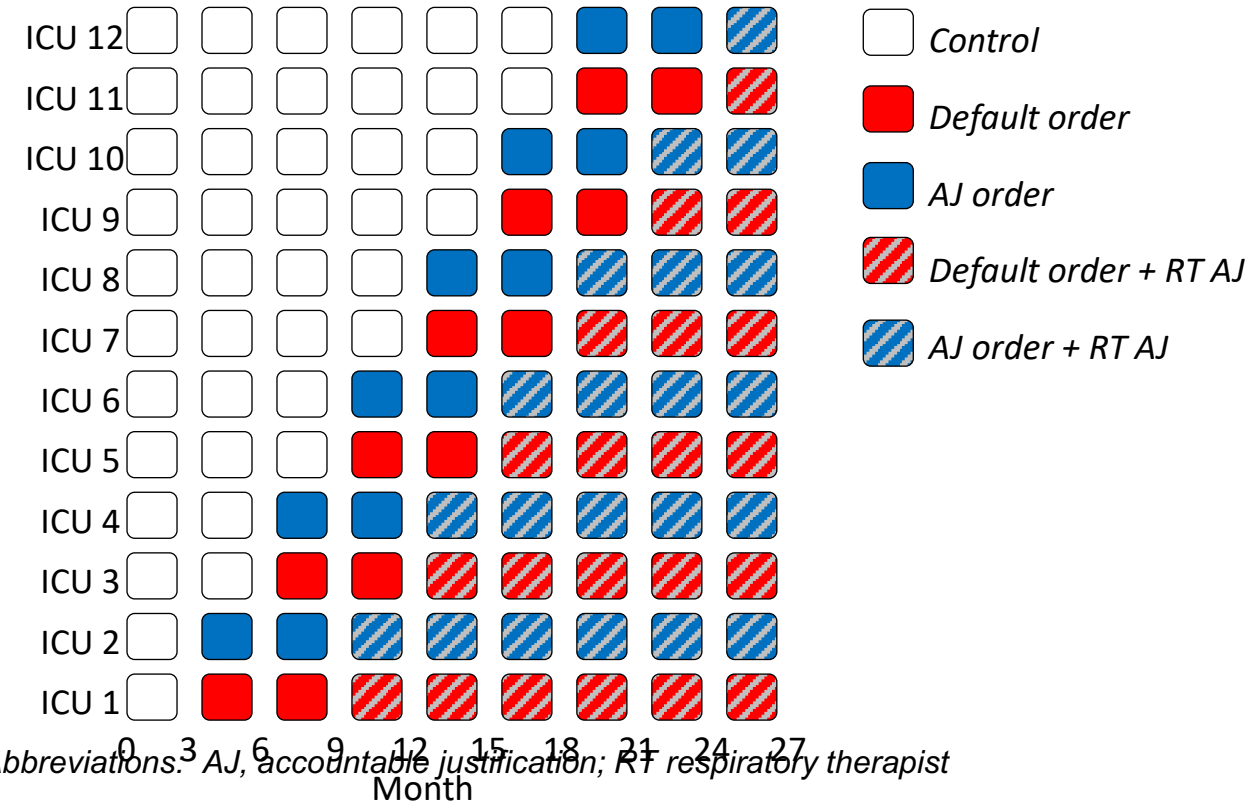
XV. References

1. Lin, Winston, Scott D Halpern, Meeta Prasad Kerlin, and Dylan S Small. “A ‘Placement of Death’ Approach for Studies of Treatment Effects on ICU Length of Stay.” *Statistical Methods in Medical* Research 26, no. 1 (February 2017): 292–311. doi:10.1177/0962280214545121.

2. Harhay MO, Ratcliffe SJ, Small DS, Suttner LH, Crowther MJ, Halpern SD. Measuring and Analyzing Length of Stay in Critical Care Trials. Med Care 2019;57:e53-e9.

Appendix A

Schematic diagram for INPUT Trial



Abbreviations: AJ, accountable justification; RT respiratory therapist

The schematic diagram above illustrates the roll-out of three implementation strategies in the stepped-wedge design: (1) a default order panel, (2) an accountable justification order panel, and (3) an accountable justification flowsheet strategy that targets respiratory therapists.