Diagnostic Stewardship for Ventilator Associated Pneumonia

Description of Outcomes and Statistical Analysis Plan

NCT05989269

August 11, 2023

Specific Aim 1: In a cluster-randomized crossover trial among 6 ICUs across 3 medical centers, evaluate the impact of a VAP diagnostic stewardship intervention on antibiotic use, VAP diagnoses, and adverse events.

Hypothesis: We expect a reduction in unnecessary antibiotics for VAP and in VAP clinical diagnoses in the intervention vs. control periods across all sites, without an increase in adverse events.

Primary Outcome: The primary efficacy outcome is reduction in proportion of unnecessary antibiotic therapy for presumed VAP in the intervention (modified reporting) period compared to control (standard reporting period), among patients not meeting clinical criteria for pneumonia and with growth of one or more organism(s) not considered normal upper respiratory flora.

Antibiotic action based on index culture result will be classified as follows:

a) Not on empiric antibiotic(s), new antibiotic started based on index respiratory culture result, completed course for VAP
b) Not on empiric antibiotic(s), no new antibiotics initiated based on culture result
c) On empiric antibiotic(s), changed based on index respiratory culture, completed course for VAP
d) On empiric antibiotic(s), continued without change following index respiratory culture, completed course for VAP
e) On empiric antibiotic(s), antibiotics discontinued based on culture results
f) On empiric antibiotics(s), but this therapy was not directed at the results of the index respiratory culture, and was not directed towards VAP

For patients where culturing was deemed inappropriate, antibiotic use in categories a, c, and d will together be considered "unnecessary antibiotic therapy for presumed VAP"

Note: Multiple respiratory cultures from the same patient may be eligible for intervention i.e., modified reporting during the intervention period. However, clinical outcomes linked to an "index respiratory culture" will be assessed using a single randomly selected respiratory culture from each included patient.

Secondary outcomes are as follows:

1. Antibiotic consumption outcomes: The following three outcomes will be measured to understand the impact on antibiotic use in ICU patients
    a. Duration of antibiotic therapy: The duration of antibiotic therapy for possible or presumed VAP among patients who underwent intervention during the intervention period and among patients eligible for intervention in the control period will be compared. Duration of antibiotic therapy will be collected for all categories above (except b), and defined as uninterrupted days of therapy, irrespective of the number of different antibiotics, for the same episode of presumed VAP, assessed at discharge or day 28 whichever occurs first. For patients in category b, duration will be zero.
    b. Antibiotic consumption during study intervention and control periods will be collected as antibiotic days of therapy (DOTs) as defined by CDC (Centers for Disease Control) NHSN as number of days that patients receive an antibiotic agent during a calendar day. The DOT for a given patient receiving multiple antibiotics on the same calendar day will

be the sum of the DOT for each antibiotic. This will be standardized to patient census of 1,000 patient days.

i. Unit-level total antibiotic days of therapy (DOT) per 1,000 ICU patient-days: These data will be obtained from the hospital antimicrobial stewardship programs and will be compared in aggregate between the intervention and control periods for all study ICUs combined, and by hospital.

ii. Unit-level respiratory antibiotic DOT: This will be captured from the EMR by counting antibiotics that are ordered for "respiratory infection" indication starting at day 3 or later of ICU stay and expressed as days of therapy per 1,000 ICU patient-days. We will validate this outcome by chart review of a small sample of patients. These will be compared in aggregate between the intervention and control periods for all study ICUs combined, and by hospital.

2. Clinical diagnoses of VAP within 7 days after the index culture - measurement of this outcome will help understand how modification of reporting would influence the likelihood of a patient receiving a diagnosis of VAP. Because there are no consensus definitions for VAP, we will use the treating clinician's documentation of pneumonia or VAP and treatment for those conditions to capture this outcome. The clinical diagnosis of VAP is an intermediate outcome that precedes the endpoint of antibiotic use. Clinical diagnoses of tracheobronchitis within 7 days after the index culture will also be recorded.

    a. We will perform a sensitivity analysis of the impact on antibiotic use outcomes after excluding patients diagnosed with tracheobronchitis within 7 days of index culture. The hypothesis is that the intervention will have limited impact on those diagnosed and treated for ventilator-associated tracheobronchitis, and thus, we may see a larger difference between intervention and control when excluding patients with diagnoses of tracheobronchitis Ventilator-free days: Ventilator-free days is defined as the number of calendar days within 28 days after the index respiratory culture on which the patient was not mechanically ventilated. Any patient dying within 28 days of the index respiratory culture collection will be assigned zero ventilator-free days. Ventilator-free days will be compared between the intervention and control periods for all study ICUs combined, and by hospital.

3. Rate of requests for complete reports in the intervention period – these data are applicable in the intervention period only. This will be a primarily descriptive analysis to help us understand if the intervention worked as intended i.e., how frequently did clinicians still want the full report with organism identification despite the nudge to consider asymptomatic colonization.

4. Adverse event outcomes*:* Death, bacteremia, and septic shock (from any cause) within 7 days of the index culture order will be recorded as potential significant adverse events from a delayed or missed true VAP diagnosis or delay in initiation of appropriate antimicrobial therapy, similar to methods described by Daley et al.[1] Incidence of these events will be compared between intervention and control periods individually for each event, and as a combined adverse event outcome. Two-person adjudication will be used to determine whether the above events were attributed to the intervention, during the intervention period.

Statistical methods for Aim 1:

Descriptive analyses using the Wilcoxon rank-sum test for continuous variables and the chi-square or Fisher's exact tests for categorical variables will be used to compare patient-level data across each intervention arm. Analyses will be conducted on an intention-to-treat basis and will account for the cluster-randomized crossover design to ensure correct type I error rates and confidence intervals (CIs). A significance level of 0.05 will be used for all statistical comparisons.

Primary Outcome: The primary efficacy outcome is the reduction in proportion of unnecessary antibiotic therapy for presumed VAP in the intervention period i.e., modified reporting period compared to control period or standard reporting period, among patients not meeting criteria for culture order and with growth of one or more organism(s) not considered normal upper respiratory flora. Antibiotic therapy given for a culture that was collected in a patient not meeting algorithm criteria will be considered unnecessary (dichotomous outcome yes/no). For the primary outcome analysis (and all other patient-level analyses below), a patient will be included once in a given admission and the outcome will be assessed on the basis of a randomly selected qualifying respiratory culture during that admission (termed index respiratory culture). We will estimate the proportion of the primary outcome (i.e., unnecessary antibiotic therapy) in each intervention arm and will compare the two groups using risk difference from a generalized linear mixed model with an identity link function. The main analysis for the primary outcome will be intention-to-treat; a secondary per protocol (as treated) analysis will be done in which any cultures during the intervention period for which a modified report was requested but either the modification was not made, or the report was modified but the full report was released upon provider request or other reason, will not be included in the intervention group.

Power and Sample Size: For the population of ICU patients for which ordering of a respiratory culture has been determined inappropriate and there is growth of one or more organism(s) not considered normal upper respiratory flora, we test the effect of the diagnostic stewardship intervention on the proportion of antibiotic prescription in a two-period cluster-randomized cross-over design, where an ICU forms a cluster. The outcome follows a binary distribution with 50% probability to receive antibiotic prescription in the control group (pC). In a conservative estimate, we expect that the intervention reduces the probability by $\Delta = 0.15\ (= \mathrm{pC} - \mathrm{pI})$ percentage points. Assuming the within-cluster within-period correlation is $\rho = 0.06$ and the within-cluster between-period correlation is about $\eta = 0.054$ [2,3] a total sample size of $N = 350$ in 6 ICUs across both periods is needed to detect an effect of $\Delta = 0.20$ with 85% power, using a generalized linear mixed model with random cluster, period, and cluster-period interaction effect. The computed total N is equivalent to a mean of about 5 patients per ICU per month, which is feasible in our experience. Given the uncertainty in predicting the actual number of patients that will be eligible to receive the intervention per ICU during the fixed 6-month period, we calculated the power for varying range of effect sizes and total N (Table).

Option	Power	Proportion control	Proportion intervention	N patients eligible for intervention per ICU per period*	Within-cluster within-period rho	Within-cluster between-period rho	Total N
1.	90%	50%	35%	30	0.06	0.054	619
2.	85%	50%	35%	30	0.06	0.054	547
3.	80%	50%	35%	30	0.06	0.054	493

4.	90%	50%	30%	30	0.06	0.054	390
5.	85%	50%	30%	30	0.06	0.054	350
6.	80%	50%	30%	30	0.06	0.054	322
7.	90%	50%	25%	30	0.06	0.054	285
8.	85%	50%	25%	30	0.06	0.054	261
9.	80%	50%	25%	30	0.06	0.054	243

Cluster- and participant-level covariates at baseline will be reported within each intervention arm. Adjusted analysis of the primary outcome will be performed as needed to control for differences in possible influential factors between intervention arms at baseline.

Statistical Methods for Secondary outcomes:
Generalized linear mixed models with a random cluster, period, and a cluster-period interaction modeling a negative binomial distribution will be used to compare the antibiotic duration count outcomes assessed at the participant-level (i.e., duration of therapy) between the intervention vs the control ICUs. The Wilcoxon rank-sum will be used to compare differences in antibiotic duration outcomes assessed at the ICU-level (total and respiratory antibiotic DOT) across intervention arm. Similar to the primary outcome model, generalized linear mixed model with random cluster, period, and cluster-period intervention effect will be used to compare other binomial participant-level outcomes (i.e., adverse events) across intervention arm. No adjustment for multiple comparisons will be made for secondary outcomes, therefore results should be considered exploratory.

Specific Aim 2: Evaluate overall impact of intervention including clinical and antibiotic outcomes using the "Desirability of Outcome Ranking (DOOR)/ Response Adjusted for Duration of Antibiotic Risk (RADAR)" methodology.

Aim 2a: Apply DOOR/RADAR methodology to measure outcomes of specific Aim 1
Hypothesis: We expect better overall patient outcomes (better DOOR ranking, accounting for duration of antibiotic use) in the intervention vs. control period.
Strategy: We will determine the probability that a patient who received the intervention has a better outcome than a participant who received the standard treatment. The clinical outcome of a participant will be ranked according to the table, from most (lower rank) to least (higher rank) desirable outcome, to obtain a 5-level ordinal outcome. We will conduct pairwise comparisons of each intervention patient to each control patient using two rules: When comparing patients with different overall clinical outcomes, the patient with a better overall clinical outcome based on the ranks in the below Table (Table 2) is determined to have the better outcome. When comparing two patients who have the same rank based on overall clinical outcome, the patient with a shorter course of antibiotics has the better outcome. Therefore, clinical outcome outranks the duration of antibiotic use i.e., a patient with a worse clinical outcome cannot have a better rank than a patient with a better clinical outcome, regardless of the duration of antibiotics.[4] The duration of antibiotic therapy for possible or presumed VAP in relationship to the index respiratory culture (see above) will be used as the antibiotic duration measure for this analysis. Note that some patients will be on antibiotic therapy for reasons other than VAP - category f above. Thus, our antibiotic use duration may be more conservative for RADAR than studies that only look at antibiotics for the specific clinical entity, but this strategy helps to avoid overestimating

the impact on ICU antibiotic use in general. We will plan to conduct a sensitivity analysis with duration of antibiotic use defined with and without category f.

Table 2. Proposed DOOR criteria rank ordered by desirability of outcome

Rank	Objective clinical outcome (must meet each criterion listed for rank)
1	Discharged home by day 14, and not readmitted between days 15-28
2	Discharged home by day 14 but readmitted between days 15-28 Discharged to facility without ventilator by day 14, and not readmitted between days 15- 28 Off ventilator by day 14 and discharged home between days 15-28
3	Discharged to facility on ventilator by day 14 and not readmitted between days 15-28 Discharged to facility without ventilator by day 14 but readmitted between days 15-28 Off ventilator by day 14 and still in hospital between days 15- 28 Off ventilator by day 14 and discharged to facility between days 15-28 On ventilator > day 14 and discharged home between days 15-28
4	Discharged to facility on ventilator by day 14 and readmitted between days 15-28 On ventilator > day 14 and discharged to facility between days 15-28 On ventilator > day 14 and still in hospital between days 15- 28
5	Died in-hospital or designated for hospice care by day 28

Statistical methods for Aim 2a*:* The DOOR/RADAR probability represents the probability that a randomly selected patient who received the intervention has a better outcome than a randomly selected participant who received the standard reporting (control) accounting for antibiotic length of therapy. For each participant in an intervention cluster (ICU), the number of participants with a less desirable DOOR rank than that participant from the same ICU but during the control period will be determined. Results are then summarized across all 6 ICUs and the DOOR probability will be computed. Inverse probability weighting methodology using patient clinical characteristics may be applied to account for potential confounding.

In addition to the DOOR analyses just described, a two-dimensional analysis of the DOOR probability and difference in the mean duration of antibiotic use in the intervention and control groups will also be performed. Only subjects with non-missing DOOR and number of days on antibiotics will be included in these analyses. The horizontal axis will be the DOOR probability (probability of a more desirable outcome when assigned to intervention vs control) based on the 5-level ordinal DOOR outcome ignoring the duration of antibiotic use. The vertical axis is the difference in the means of the observed duration of antibiotic use (intervention minus control). A result in the lower right quadrant represents more desirable results for standard reporting (control), while a result in the upper left represents more

desirable for modified reporting (intervention). The other two quadrants represent tradeoffs for clinical outcomes and antibiotic use.

Aim 2b: Compare results of standard outcome analysis to those of DOOR analysis.

Strategy: We will conduct descriptive qualitative and quantitative comparisons of the results of Aim 1 and Aim 2a with respect to both statistical and clinical significance. *Hypothesis:* Compared to standard metrics used in infectious diseases research, which separates efficacy and safety into different measures, the more patient-centered DOOR measurement will better differentiate risk versus benefit of the intervention relative to the control groups, by considering any reductions in potentially unnecessary antibiotic therapy in the context of the overall clinical outcomes in a single measure.

Interim analyses

We will monitor fidelity of the intervention via monthly review of the data and calls with the study sites. This will include evaluation of the numbers of cultures evaluated, number (and proportion) eligible for intervention, number (and proportion) of cultures intervened upon, and number (and proportion) of requests to release complete reports.

No interim efficacy analyses are planned. No interim safety analyses are planned for this minimal-risk pragmatic cluster trial and therefore, no stopping rules have been developed. The outcomes analyses specified above will not occur until all sites have completed the planned intervention period.

REFERENCES

1. Daley P, Garcia D, Inayatullah R, Penney C, Boyd S. Modified Reporting of Positive Urine Cultures to Reduce Inappropriate Treatment of Asymptomatic Bacteriuria Among Nonpregnant, Noncatheterized Inpatients: A Randomized Controlled Trial. *Infect Control Hosp Epidemiol*. 2018;39(7):814-819. doi:10.1017/ice.2018.100
2. Blanco N, Harris AD, Magder LS, et al. Sample Size Estimates for Cluster-Randomized Trials in Hospital Infection Control and Antimicrobial Stewardship. JAMA Netw Open. 2019;2(10):e1912644. doi:10.1001/jamanetworkopen.2019.12644
3. Arnup SJ, McKenzie JE, Hemming K, Pilcher D, Forbes AB. Understanding the cluster randomised crossover design: a graphical illustration of the components of variation and a sample size tutorial. Trials. 2017;18(1):381. doi:10.1186/s13063-017-2113-2
4. Evans SR, Rubin D, Follmann D, et al. Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). *Clin Infect Dis*. 2015;61(5):800-806. doi:10.1093/cid/civ495