STATISTICAL ANALYSIS PLAN

We will describe continuous variables using mean (standard deviation) or median (interquartile range), as appropriate. Categorical variables will be described using frequency (percent). Unless otherwise specified, hypothesis tests will be two-sided with results considered statistically significant at $p<0.05$. Analysis will be performed on an intention-to-treat basis. In all cases, the unit-of-analysis will be an individual patient. Analysis will be performed in SAS.

Primary Analysis Cohort

The primary analysis will be an intention-to-treat analysis among the patients who were at risk of bundle non-compliance: i.e., those for whom a reminder page was generated.

Primary Outcome

The primary outcome will be the difference in overall ordering compliance between patients randomized to intervention vs. usual care. Overall compliance will defined as orders within the time limit for antibiotics, blood culture collection, and initial lactate measurement within 3-hours and orders for a repeat lactate measurement both within 6-hours of time-zero and 3-hours of initial lactate measurement for patients whose initial lactate measurement was ≥2.0mmol/L. Study groups will be compared based on the absolute difference in percent compliance, 95% confidence intervals will be calculated for the difference, and p-values will be computed using Chi-Square test.

Sensitivity Analyses: Primary Outcome

Sensitivity analyses will compare overall bundle compliance between patients randomized to intervention vs. usual care using logistic regression. Models will adjust for day vs. night shift assignment, service designation (medicine vs. surgery), and number of days since study enrollment first began at the time of the patient's enrollment.

Additional Process and Care Delivery Outcomes

Analyses will be conducted as above for the additional process outcomes. For ordinal outcomes, (e.g., the number of bundle elements ordered with compliance), proportional odds regression adjusted as above will be used. In addition, for care delivery outcomes, all process outcomes above will be assessed (exactly as above), but using the time of care delivery rather than order time (i.e., the administration rather than order time for antibiotics, and the collection time for blood cultures and for lactate measurements.)

Exploratory Clinical and Balancing Outcome

The exploratory patient outcome endpoints will be as analyzed as below. In all cases, models will adjust for age, first SOFA score, initial lactate, and Charlson Comorbidity Index. These outcomes are exploratory and adjustments for multiple comparisons will not be made.

The main exploratory clinical outcome patient endpoint is 28-day mortality. We will compare mortality between groups using logistic regression. Additional binary endpoints that will analyzed analogously are mechanical ventilation or death within 72-hours, and ICU admission or death within 72-hours, mechanical ventilation or death during hospitalization, and ICU admission or death during hospitalization. Additionally, the same method will be used to analyze antibiotic discontinuation within 48-hours and culture negativity outcomes.

Hospital length of stay will be analyzed using a Fine-Gray competing risk regression with informative censoring for hospital mortality. Discharge alive will be considered the event of interest and hospital mortality will be considered a competing event.

Sensitivity Analyses: Overall Cohort

The above analyses will be recapitulated as an additional sensitivity analysis among all enrolled patients, including those for whom no alert was generated. This will be performed based on the rationale that paging alerts for one patient could potentially lead to contamination bias if they caused a clinician to act on other patients under their care at the time for whom they did not receive a paging alert, or maturation bias if the alerts altered their care of future patients if treated during period of allocation to the control arm that occurred after a period of being allocated to the intervention arm.

Prespecified Subgroup Analyses

We will include four several pre-specified subgroup analyses.

1) Patients who had a reminder page generated specifically for antibiotics versus those who had pages generated for bundle elements other than antibiotics.
2) Patients who did versus did not test positive for COVID-19.
3) Patients who had an initial lactate ≥ 4.0 mmol/L vs. 2.1-3.9 mmol/L vs. < 2.0 mmol/L.
4) Patients who with versus without hemodynamic instability within 1 hour of time-zero. Hemodynamic instability will be defined as (systolic BP < 90 OR diastolic BP < 50 OR mean arterial pressure < 65 mmHg) AND (heart rate > 100 bpm).

Interim Analysis

An interim analysis with a non-binding design (i.e., no stopping for futility) will be conducted to assess for a major safety signal. This analysis will occur when total enrollment reaches n=1,170 patients. The trial will be stopped if $p < 0.0056$ for the mortality difference between randomized groups among the patients who triggered a page alert.

This criterion is based on a minimum important absolute difference of 8% with $\alpha=0.05$, $\beta=0.2$, and control group mortality = 20%. A one-sided hypothesis test will be used at this stage in order to reduce the sample size needed to detect a safety signal as it will not be assumed that the intervention can plausibly increase mortality. Assuming an interim analysis occurring at 25% of the total enrollment needed for a trial powered to detect this effect, the estimated enrollment target would be n=514 total patients. Additionally, it will be assumed that any mortality effect would be driven by the population for whom practice is changed. Based on pilot data, this is assumed to be 25% of patients for whom an alert page is sent. Further, it is assumed from pilot data that only 33% of patients who trigger the Sepsis BPA will subsequently trigger the reminder page. Thus, an appropriately powered trial would require an estimated n=6,168 total patients to identify an effect within this subset, and our interim analysis will therefore be performed when total enrollment reaches n=1,170, assuming there will be 390 patients for whom an alert is triggered, and 130 patients for whom practice will change. The trial will be stopped at this time if, among the patients who triggered a page alert, the mortality difference between randomization arms corresponds to a one-sided $p < 0.0056$, which is estimated to correspond to a $p < 0.0005$ among the patients for whom paging changed practice, a group that is not directly identifiable.