Category :**Sepsis: biomarkers**

A263 - ***Computational endotypes of sepsis and septic shock derived from time series data***

J Guo [1] ; H Kim [2] ; R Stevens [2]

[1]Johns Hopkins University, Baltimore, United States, [2]Johns Hopkins School of Medicine, Baltimore, United States

Introduction:

While sepsis has been characterized through the sepsis-3 consensus criteria as organ dysfunction and dysregulated host response to infection, there remains significant heterogeneity in the prediction, identification, treatment, and outcomes of sepsis and septic shock. This project aims to discover novel subphenotypes of sepsis and septic shock using highly granular time-series data from electronic health records.

Methods:

From the Philips multi-center eICU database, we calculated Sequential Organ Failure Assessment (SOFA) scores and determined infection status by diagnosis and/or positive culture for each adult patient every 5 minutes. Based on the sepsis-3 criteria met at a timepoint, patients were categorized into four data-driven cohorts, or "statuses" (No Sepsis, Suspected Sepsis [only organ dysfunction], Sepsis, or Septic Shock). 24 clinical and physiological features from 24 to 48 h after ICU admission were used for unsupervised clustering algorithms. Endotypes from clustering were characterized by length of stays (LOS) and hospital discharge outcomes.

Results:

Clustering analysis identified five distinct groups. Cluster 1 contained a large proportion of suspected sepsis patients and 3 contained most septic patients. Clusters 2 and 4 contained the majority of septic shock patients, with 2 having the larger proportion and the overall longest hospital LOS (10.7 days) and largest mortality (35%). Cluster 5 contained nearly all non-septic patients. Clusters with mixes of statuses highlight potential similarities among patients with different sepsis-3 classifications, and the limitations of consensus criteria.

Conclusion:

Using temporal data with a broader range of features, we identified five distinct endotypes for sepsis and septic shock. The high-frequency time series may allow for a more refined, data-driven method to characterize the evolution of sepsis and stratify patients that may be misidentified through sepsis-3, better informing prediction, detection, and treatment.

Image :

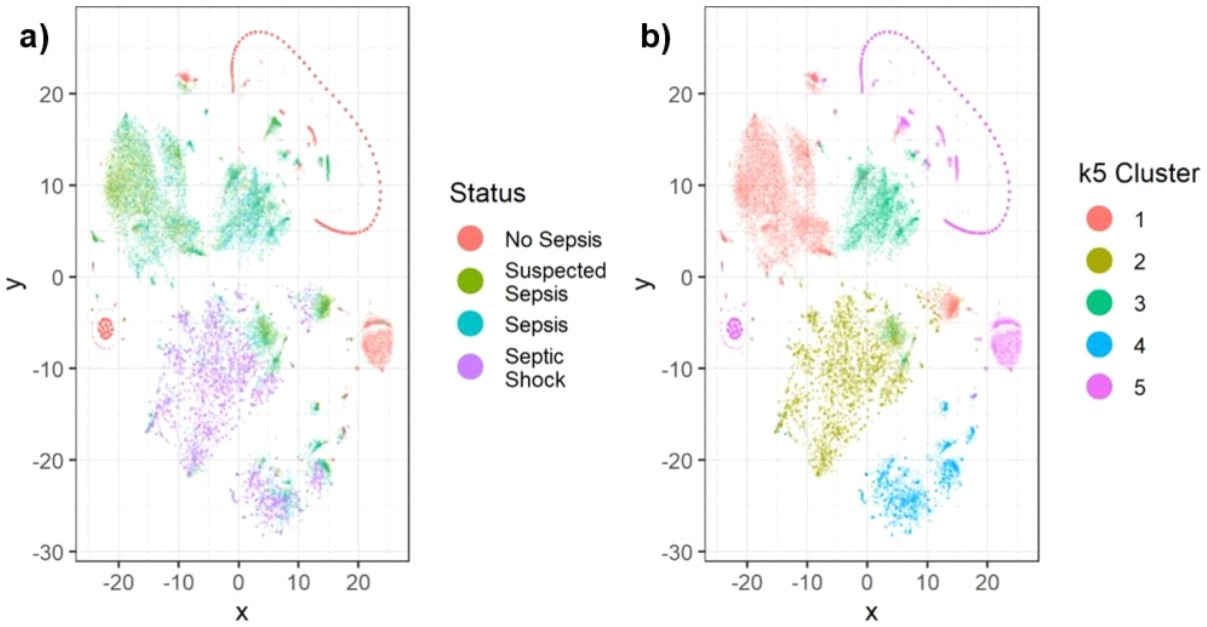

Figure 1. Identification of Computational Endotypes. a) Visualization of patient cohorts, or "statuses," with t-

Distributed Stochastic Neighbor Embedding (tSNE). The scaled dataset containing 24 clinical and physiological variables was visualized with tSNE, a dimensionality reduction algorithm that probabilistically maps complex high-dimensional data (i.e. 24-dimensional data) to lower dimensions and identifies similar points. Each point represents one 5 min timepoint sample for each patient from 24-48h after ICU admission. 13,487 timepoints (80% of septic shock timepoints) were randomly sampled for each status. b) Visualization of clusters from K-Means (k5) on scaled data overlapped on previous tSNE plot.