



**UNDERSTANDING THE CHANGES OF PATIENTS IN THE INTENSIVE CARE UNIT:
A CASE STUDY OF SEPSIS AND ACUTE KIDNEY INJURY**

Kyle Kazemini, Samir Abdelrahman, MSc, PhD
Department of Biomedical Informatics
University of Utah School of Medicine

Abstract

Sepsis is a dysregulated host response to an infection. This response may include tissue hypoperfusion, organ dysfunction, hypotension, hypoxia, and organ failure. Acute Kidney Injury (AKI) is defined as an increase in serum creatinine levels by at least 0.3 mg/dl or to 1.5 times baseline.

Development of AKI during sepsis increases patient morbidity, predicts higher mortality, has a significant effect on multiple organ functions, is associated with an increased length of stay in the intensive care unit, and hence consumes considerable healthcare resources. Although there are treatments for both conditions, it is imperative that clinicians start treatment early. We use clustering analysis in hopes of understanding the similarities and dissimilarities in sepsis and AKI patients in order to inform clinical interventions.

For our analysis, we use the Medical Information Mart for Intensive Care (MIMIC)-IV database, which provides critical care data for over 40,000 patients admitted to intensive care units. We use MySQL to select 51 hourly time steps for ICU patients, which include patient information, vital signs, and sepsis and AKI outcomes. This data can then be transformed into a time series dataset.

We use a Python library called tslearn for machine learning on time series data. We choose the KShape clustering algorithm for its accuracy and efficiency. The KShape clustering algorithm will allow us to evaluate similarities and dissimilarities in sepsis and AKI patients using the clusters themselves and other powerful evaluation tools such as silhouette scores.

We found silhouette scores of -0.0213 and 0.19642 for sepsis and AKI clusters respectively. In short, this means that we have found similarities, but there is still work to be done to improve the quality of the clusters. We are considering adjusting the number of time steps, choosing different features to evaluate, and using different clustering algorithms such as KMeans, DBSCAN, OPTICS, BIRCH, or hierarchical clustering.

This research was supported by a Supplement to the NLM Training grant T15 LM00712418. Computer resources were provided by the University of Utah Center for High Performance Computing, which has been partially funded by the NIH Shared Instrumentation Grant 1S10OD02164401A1.