Iowa Initiative for Artificial Intelligence
Final Report

Project title:	Novel Machine Learning Algorithms to Risk Stratify Patients with Syncope Presenting to the Emergency Department	
Principal Investigator:	Milena A. Gebska , MS, Brian Olshansky , Sangil Lee , Deepak Kumar Pasupula ,	
Prepared by (IIAI):	Avinash Mudireddy	
Other investigators:	Antony Anandaraj , Samuel Johnston , Tyler Bullis , Aron Evans , Giselle M.Statz , Jon N. Van Heukelom	
Date:	22/10/2022	
Were specific aims fulfilled:	Yes	
Readiness for extramural proposal?	No	
If yes … Planned submission date		
Funding agency		
Grant mechanism		
If no … Why not? What went wrong?	The team must assure availability of relevant outcome data and identify several main specific aims for the follow-up grant submission.	

Brief summary of accomplished results:

Research report:
Aims (provided by PI):

Syncope, one of the most common conditions seen in medical practice, is a major cause for emergency department (ED) visits and subsequent hospitalization. Some etiologies for syncope are easily identified and benign, but most are difficult to determine. Often extensive, expensive, unnecessary, and potentially harmful evaluations are undertaken due to fear of missing life-threatening cardiogenic or otherwise serious causes.1-4 Cardiogenic causes can be difficult to identify in the ED. Thus, patients are often admitted for further observation and workup. Creating an effective point-of-care risk stratification tool could help ED physicians and other first responders identify patients at high risk for sudden cardiac death and other adverse outcomes. This tool could have the potential to save lives, reduce risk of syncope recurrence, prevent unnecessary hospitalizations and testing, and save money.

Aim 1. To create an exploratory model with AI/ML using retrospective EMR data-driven approaches to risk-stratify patients presenting with syncope into *low-risk, intermediate-risk,* and *high-risk* groups, (specifically regarding cardiogenic syncope) to better define which patients may be safely discharged from the emergency department and which patients would warrant further cardiac workup and/or hospital admission.

Aim 2a. To create a point-of-care risk stratification tool based on patterns identified in the AI/ML model.
Aim 2b. To implement this tool into the UIHC Emergency Department workflow and validate its effectiveness.

Data for Aims:
We have access to the Nationwide Emergency Department Sample (NEDS) and plan to utilize it for the Pilot project. The NEDS is the largest all-payer national ED visit database in the US, sponsored by the Healthcare Cost and Utilization Project (HCUP), and yields estimates for hospital-based ED visits. The NEDS is constructed annually by sampling procedure and diagnostic codes from the State Emergency Department database and State Inpatient Database. The thirty-seven geographically dispersed states capture 68.7% of the total population and

78.2% of all ED visits in the US. The NEDS provides de-identified patient demographic and clinical variables (e.g. age, sex, and comorbidities), variables to assist with a national estimate (discharge weight, stratum used for weighing), outcomes during ED visits (e.g. discharge to home, admission to hospital, or death during visit), and disposition after subsequent inpatient care for those admitted to the hospital.

Using the NEDS, we aim to access patient data from 2006 to 2019 to help identify all patients presenting to the ED with syncope as their primary diagnosis using ICD-9 and ICD-10 codes for syncope and collapse (780.2 and R55 respectively) alone, or in conjunction, with other medical conditions (refer to Appendage 1.0 for details). Subsequently, we will analyze patients who were admitted versus those who were discharged from the ED and, using input variables listed below (X), we expect to detect patterns that can stratify syncope patients into *low risk* (safe for ED discharge), *high risk* (warranting imaging, procedures, and hospital admission) and *intermediate risk* (unclear disposition) (Y).

- **Input Variables (X)** – Patient demographics, Patient vital signs, Past Medical History and Specific symptoms through ICD codes
- **Outcomes (Y)** – Low risk (Discharged from ED), Intermediate risk (Admitted and had negative workup), High risk (Admitted and required invasive procedures)

Issues with provided Aims and remodified Aims by consulting IIAI:

Stratifying the patients into low, high and intermediate risk was not feasible due to the lack of labeled data to train the model. Hence, classifying the patient into the mentioned categories was not possible.
We tried to treat the aim as a clustering problem. However, the definitions to finally label the clusters into low, high, and intermediate risk could not be established through the results. The observed characteristics of each cluster were indistinguishable.

Later, the team settled to work on the following Remodified aim:
Predict Length of Stay for Patients Admitted with Syncope from the Emergency Department.

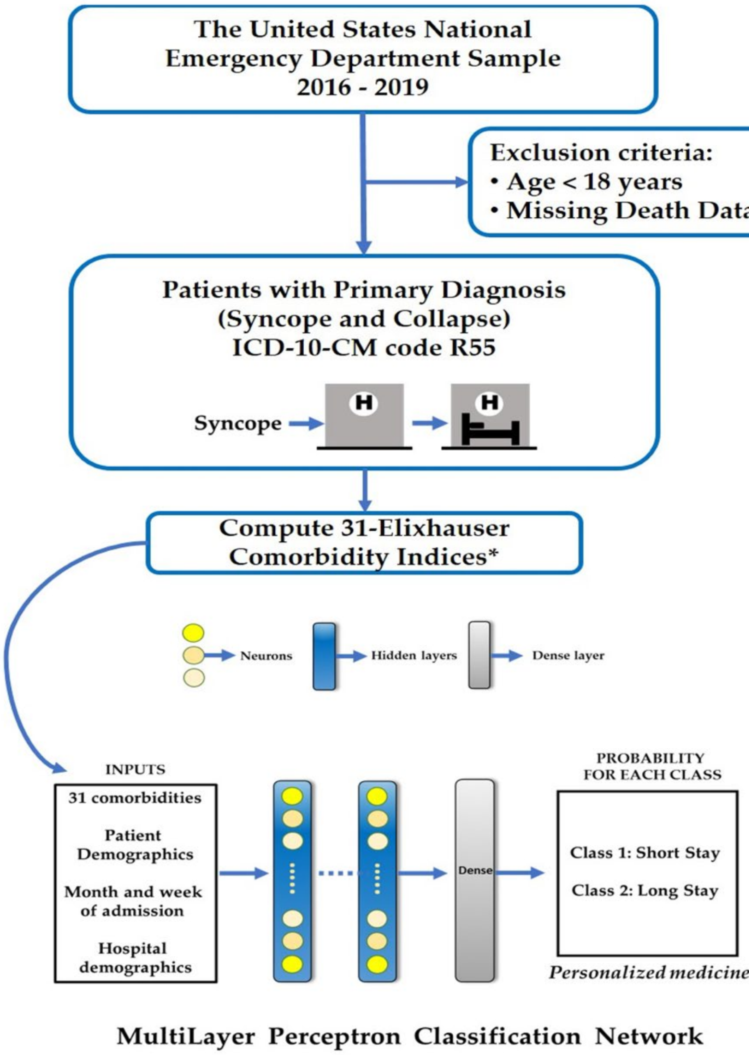
AI/ML Approach:

NEDS data(2016-2019) is filtered to exclude the population sample of Age<18 and those without mortality data. The NEDS data contains the ICD-10-CM diagnosis codes. However, for our aim, these codes are used to compute 31 Elixhauser comorbidity indices (ECI) to utilize them in the ML algorithm as a representation of personalized cardiovascular risk factors. ECI is a validated method for categorizing patient-specific comorbidities in a large administrative database based on ICD diagnosis codes. Also, we categorized LoS into short stay (negative class) and long stay (positive class). The modified input and output variables are:

- **Input Variables (X)** – Patient demographics, Month and week of admission, Rurality based on metropolitan statistical area (MSA) status, whether the ED is affiliated with an academic institution or community hospital, and 31 Elixhauser comorbidity indices.
- **Outcomes (Y)** – short stay vs Long stay

We proposed to use the Multilayer perceptron (MLP) version of Artificial Neural Network (ANN) to solve this classification-based prediction problem.

In this classification problem, the measure of the performance of the network is calculated through indicators such as Area Under Receiver Operator Characteristics – AUC, Precision, Recall, F1, and Average Accuracy.
The approach is depicted in the below image.



Experimental methods, validation approach:

As a part of the experimental setup, five models considering different short/long separation cutoffs of LoS were created. Initially, LOS is a continuous variable. However, for clinical relevance, LOS is converted to a categorical variable by thresholding. The resultant model targets are:

- short versus long stay models ≤0 days (indicating ED discharge without hospitalization) versus >0 days,
- ≤24 hours versus >24 hours
- ≤48 hours versus >48 hours
- ≤4 days versus >4 days, and
- ≤7 days versus > 7 days.

Moreover, in-hospital mortality was considered a competing outcome for hospital LoS and it was categorized as a Long stay.

As a preprocessing step, the input variables are encoded as 72 dummy variables as all of them are categorical variables. We also observed that the Target class (LOS) is highly imbalanced. Hence, we upsampled the data set to balance both classes to be of equal proportions.

The train: validation: test split is 64:16:20. That is, the test set is 20% and the remaining data is further split into 80:20 proportions to form the train and validation set.

Model design:

Our neural network has three hidden layers with 64, 32, and 16 neurons fired by rectified linear unit activation. Each layer is further treated with batch normalization and dropouts (0.05). The output layer has a binary outcome with "sigmoid" activation.
We used Adam optimizer (lr=0.001) and binary cross entropy loss as optimizer and loss function, respectively. The model is trained for 750 epochs with 8192 as batch size.

Results:

During our experiments we observed the following results for the 5 different models:

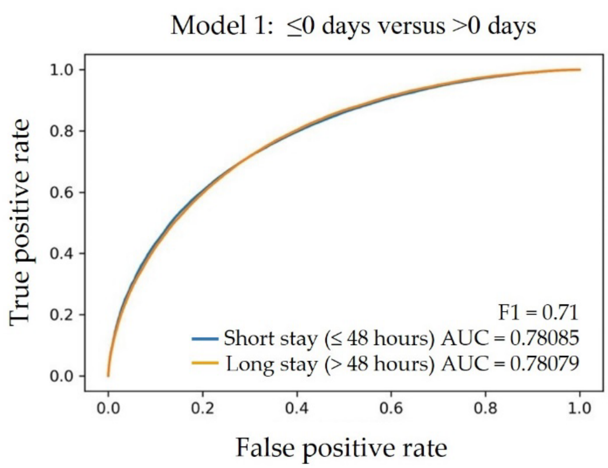
Table 1. The predictive performance of hospital length of stay

Length of Stay	AUC*	Precision	Recall	F1	Average Accuracy
≤ 0 days#	0.78	0.70	0.72	0.71	0.71
≤ 24 hours	0.79	0.72	0.72	0.72	0.72
≤ 48 hours	0.81	0.72	0.76	0.74	0.73
≤ 4 days	0.84	0.76	0.75	0.76	0.76
≤ 7 days	0.88	0.78	0.83	0.81	0.80

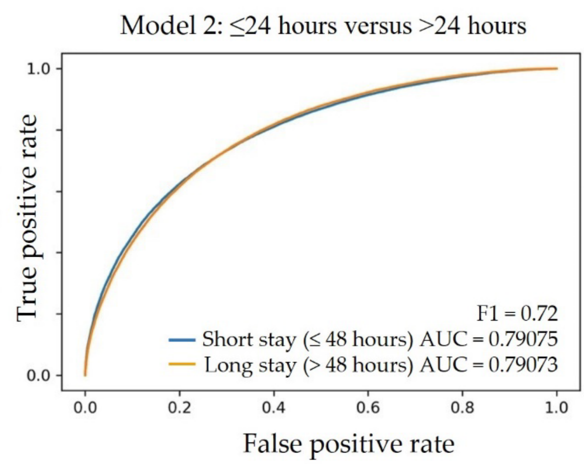
* AUC - Area Under the Curve
\# <0 days indicates discharge from ED

To be clinical relevant, the LoS cut-off value of ≤48 hours versus > 48 hours was considered to be the most practical in the real-world setting.

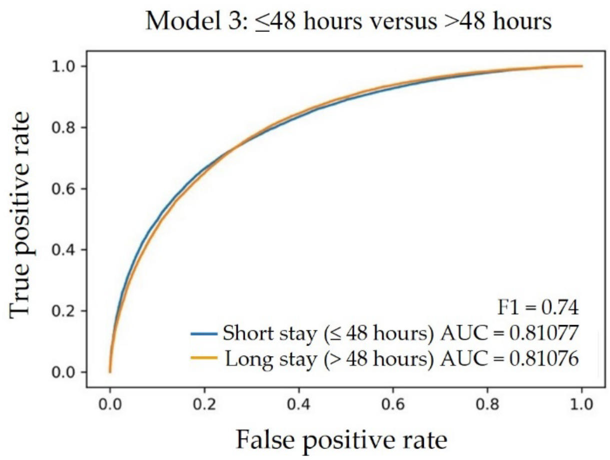
The plots for the respective ROC curves are as below.



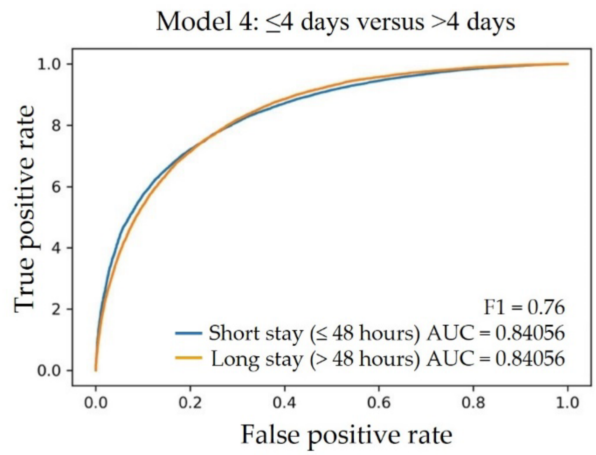
(**a**)



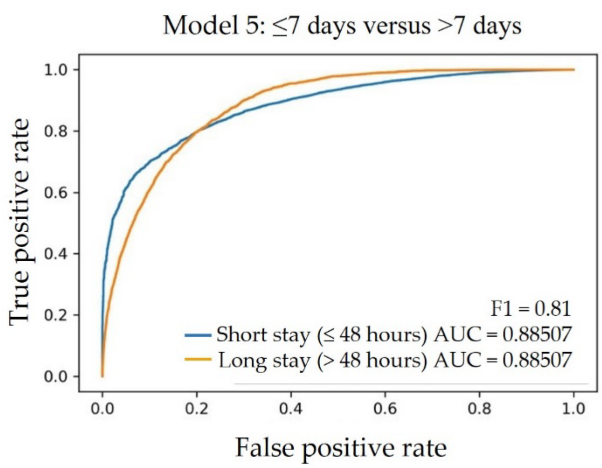
(**b**)



(c)



(d)



(e)

Other Results:

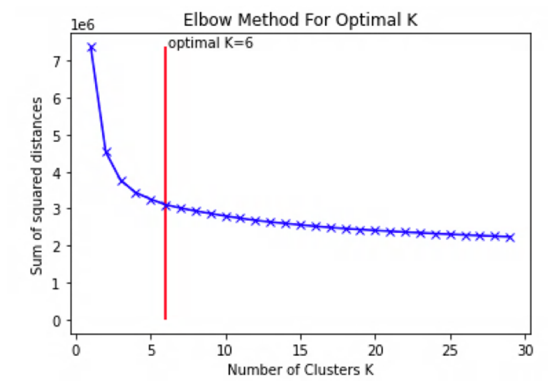
<u>Clustering Problem:</u>

As mentioned in the remodified Aim, we also worked on clustering the NEDs patient data using Patient demographics, Month and week of admission, Rurality based on metropolitan statistical area (MSA) status, whether the ED is affiliated with an academic institution or community hospital, ICD diagnostic codes, CPT codes, and PR codes.

An Autoencoder network is used to reduce the dimensionality of the data. Each layer in the model contained 1024, 512, 256, 16 (Encoder) and 16, 256, 512, 1024(Decoder) neurons. RMSProp is used as optimized and binary cross entropy is used as the loss function because all the values are categorical.

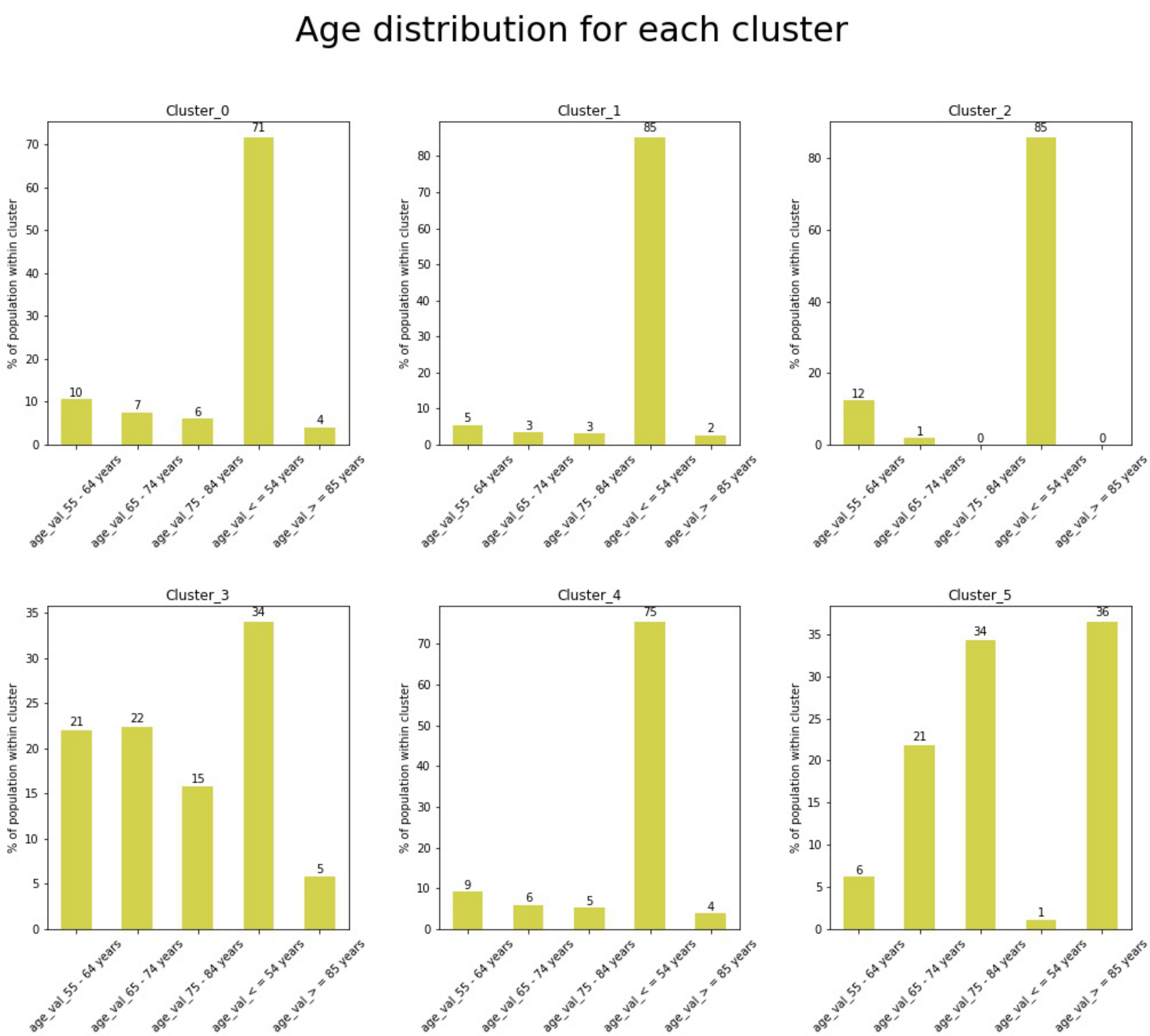
We could compress the dimensions to 16 with an accuracy of 99.84%.

The compressed dimensions are used to build a K-means clustering algorithm. The results were as shown in the picture below.

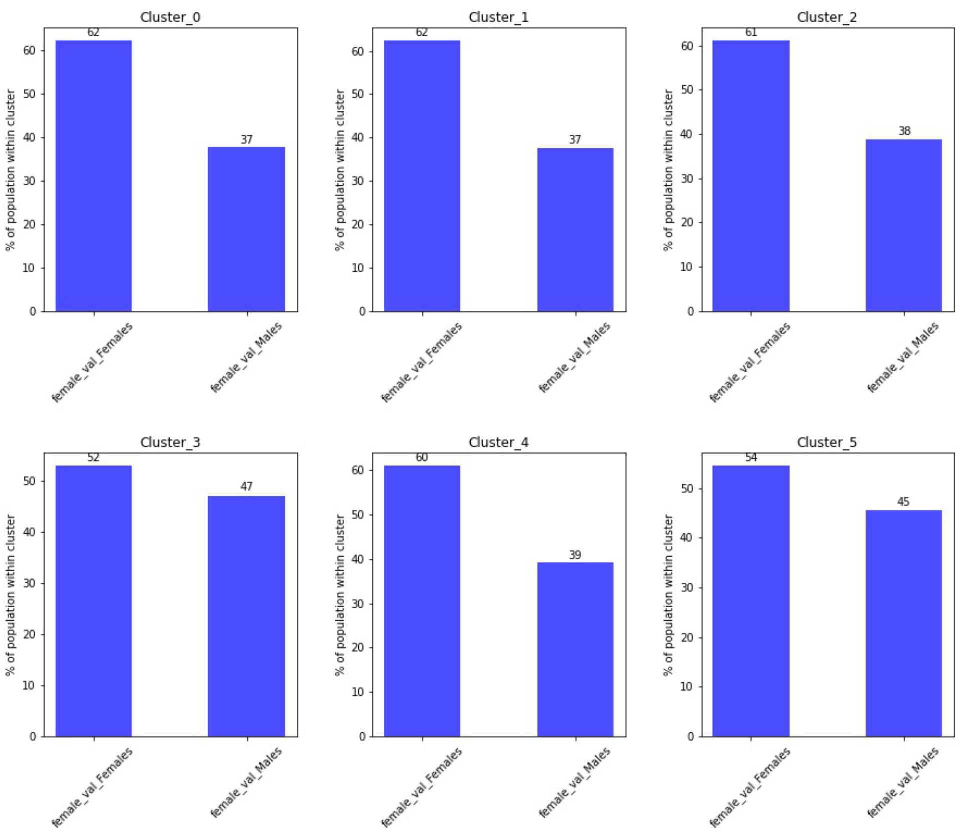


We chose the optimal cluster size as 6.

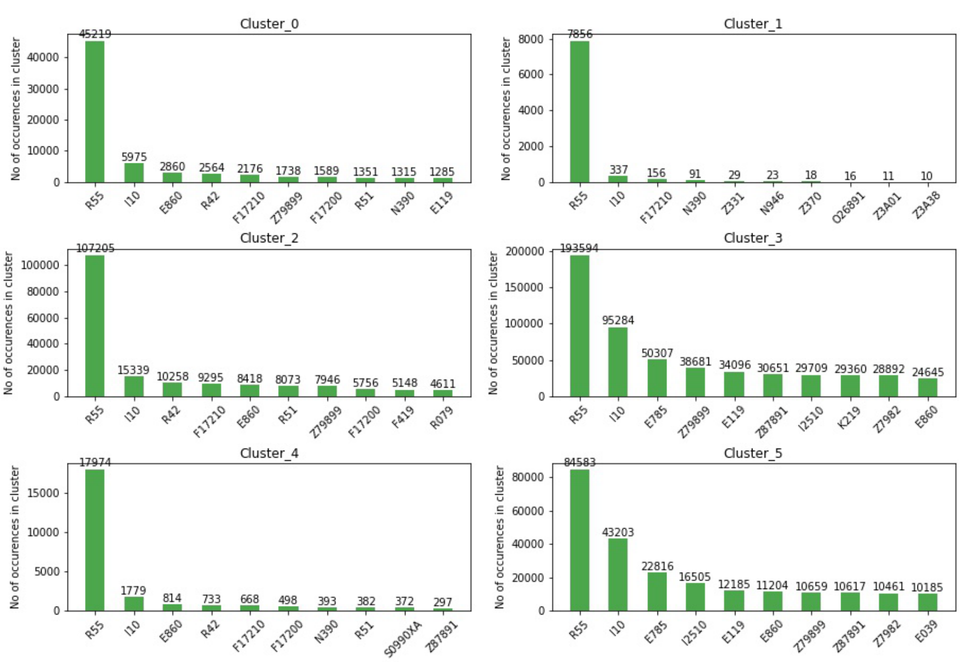
On these clusters we performed a statistical analysis to define a mechanism to label the clusters. The results are as follows.



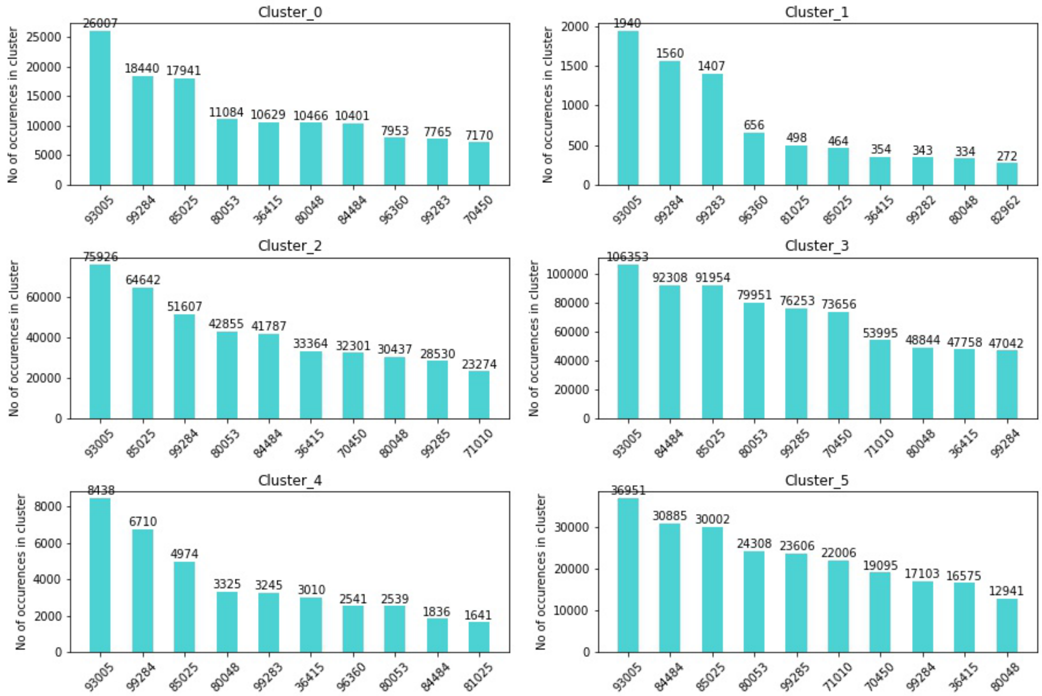
Gender distribution for each cluster



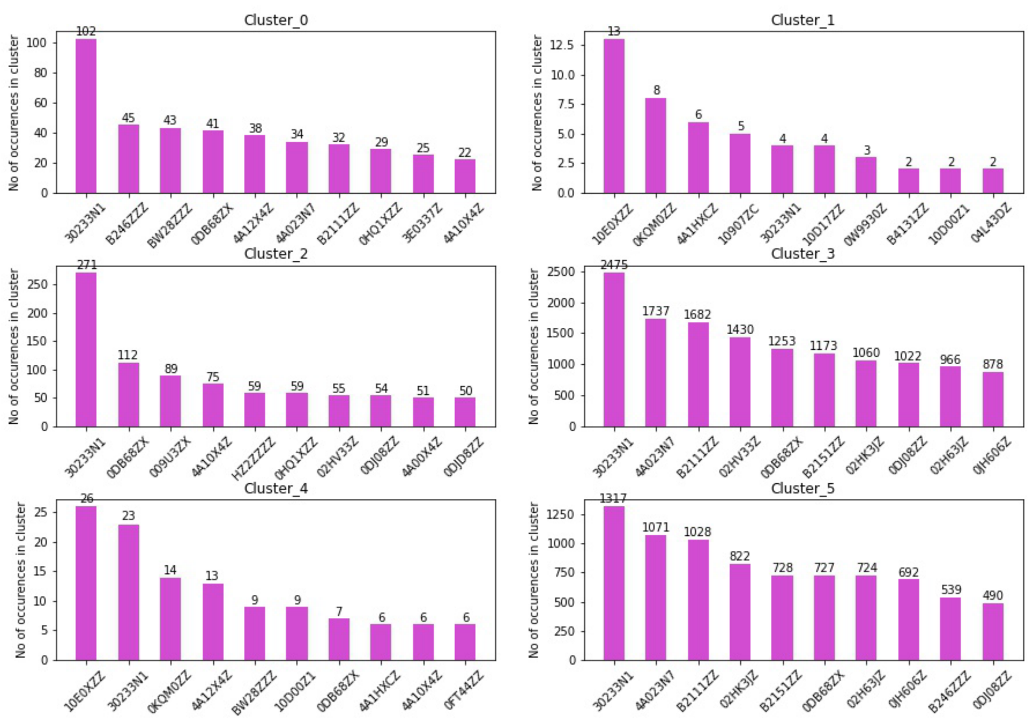
Top 10 ICD codes per cluster



Top 10 CPT codes per cluster



Top 10 PR codes per cluster



The observation is that the distinction among the culsters is not identified as a definitive. More or less all the clusters seems to overlap.
Hence, we concluded not to purse clustering.

Project Outputs – papers submitted/published

- 1 ACC/23 abstract submitted:
    - Giselle M. Statz, Sangil Lee, Avinash Reddy Mudireddy, Deepak Kumar Pasupula, Mehul Adhaduk, E. John Barsotti, Milan Sonka, Tyler Bullis, Samuel Johnston, Aron Z. Evans, Brian Olshansky, and Milena A. Gebska: MACHINE LEARNING CAN HELP PHYSICIANS PREDICT LENGTH OF STAY AFTER SYNCOPE
- 2 journal submissions:
    - Sangil Lee, Avinash Reddy Mudireddy, Deepak Kumar Pasupula, Mehul Adhaduk, E. John Barsotti, Milan Sonka, Giselle M. Statz, Tyler Bullis, Samuel L. Johnston, Aron Z. Evans, Brian Olshansky and Milena A. Gebska: Novel Machine Learning Approach to Predict and Personalize Length of Stay for Patients Admitted with Syncope from the Emergency Department, submitted to Journal of Personalized Medicine, submitted in September 2022.
    - Aron Z. Evans, Giselle M. Statz, Mehul Adhaduk, Tyler C. Bullis, Milena A. Gebska, Samuel L. Johnston, , Sangil Lee, Jon Van Heukelom, , Deepak K. Pasupula, , Milan Sonka, Avinash R. Mudireddy, Franca Dipaola, , Fabrizio Ricci, Madeleine Johansson, Robert S. Sheldon, Venkatesh Thiruganasambandamoorthy, Rose-Anne Kenny, Brian Olshansky: Can Artificial Intelligence Improve Syncope Management? A White Paper from an International Panel, submitted to Journal of the American College of Cardiology (JACC), submitted in October 2022.