Mining Malpractice Claims to Identify Disparities between Younger and Older Adults

Wanting Cui, M.A, Joseph Finkelstein, M.D, PhD
Icahn School of Medicine at Mount Sinai, New York, NY

Introduction

Older adults are considered one of the most vulnerable population in the society. The goal of this research is to explore characteristics of malpractice claims in older adults and to use unsupervised learning to identify potential cohorts in older adults.

Methods

The National Practitioner Data Bank’s (NPDB) Public Used Data File was used to create analytical dataset for this study. Care delivery entities are required to report any malpractice payments to the data bank. All data were stratified into 3 age levels. We defined older adults as patients over 70 years old, middle-age adults as patients between 40 to 69 years old, and young adults as patients between 20 to 39 years old. We performed an exploratory data analysis and calculated summary statistics. In cluster analysis, we compared three algorithms, hierarchical clustering, K-Medoids and K-Prototype, due to the non-numerical characteristics of the dataset. Sum of squared error and silhouette value were used to evaluate the performances of these algorithms.

Results

Over 113, 000 malpractice claims in a period of 10 years were analyzed in this study. Around 12% of malpractices were related to older patients. In the past 10 years, around $32 billion went into malpractice payments out of which $2.7 billion were awarded to older adults. In addition, there were over 600 death cases of older patients each year, which constituted almost 45% of malpractice cases for older adults (Figure 1). Despite the high death rate, older adults only received a median payment of $128,866 per case of death, which was over 50% less than the amount that middle-aged and young adults received. The payment disparity persisted across different malpractice types.

In cluster analysis, we used Gower coefficients to calculate dissimilarity matrix and K-Prototype was the best model to use. 5 clusters of older patients were identified. Payment amount, patient’s gender, patient’s type, malpractice allegation group, first specific allegation, multiple allegations and patient’s residential state were the important variables to distinguish clusters.

Figure 1. Proportion of Claims by Age and Outcome

Conclusion

Older adults were paid significantly less for the same malpractices claims as compared to younger adults. Distinct clusters were discovered. The NPDB database is an important dataset to study malpractice claims in older patients. Further analysis of information extracted from this dataset and building predictive models are warranted.