Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

USING LONGITUDINAL HEALTH RECORDS TO SIMULATE THE IMPACT OF NATIONAL TREATMENT GUIDELINES FOR CARDIOVASCULAR DISEASE

Daniel F. Otero-Leon

University of Michigan
1205 Beal Ave
Ann Arbor, Michigan 48109, USA

ABSTRACT

Continuous tracking of patients' health data through electronic health records (EHRs) has created an opportunity to predict healthcare policies' long-term impacts. Despite the advances in EHRs, data may be missing or sparsely collected. This article develops a simulation model to test multiple treatment guidelines for cardiovascular disease (CVD) prevention. Our methodology uses the EM algorithm to fit sparse health data and a discrete-time Monte-Carlo simulation model to test guidelines for different patient demographics. Our results suggest that, among published guidelines, those focusing on reducing CVD risk can reduce treatment without increasing the risk of severe health outcomes.

1 INTRODUCTION

Over the last couple of decades, electronic health records (EHRs) have emerged as a standard for tracking important medical data for patients seen in health systems. As a result, health systems are amassing large amounts of longitudinal data that can be used to create a deeper understanding of how patients' health conditions change over time. This resource creates opportunities for health systems to extend beyond the patients' current health status and use simulation-based approaches to guide screening and treatment decisions. We focus on cardiovascular disease (CVD) prevention as an important example of how longitudinal EHR data can improve care. We chose CVD prevention because it is a common condition requiring treatment decisions over time. In the United States, the two most often used treatment guidelines are those developed by the American College of Cardiology (ACC) (Grundy et al. 2018; Whelton et al. 2017) and the Joint National Committee (JNC) (James et al. 2014). The variations between them lead to differences in whether and when patients start the treatment and their risk of having a CVD event.

Additionally, from a physician's point of view, the data is sparsely collected. We apply EM-algorithms to fit this behavior, applied before in many healthcare settings (Yeh et al. 2010). To the best of our knowledge, this is the first application for CVD prevention.

The objective of this study is twofold. First, we provide a case study of developing a discrete-time Monte-Carlo simulation model using CVD longitudinal EHR data from the Veteran Affairs (VA) health system. Second, we use the model to evaluate the impact of the most well-known U.S. treatment guidelines for prevention of CVD. As a variant of previous studies, we propose a framework that feeds from sparse longitudinal data to simulate patients' behavior throughout their lifetime.

2 PROPOSED METHODOLOGY

We divide our framework into two steps. 1) Estimate Markov chain parameters for cholesterol and blood pressure, and 2) simulate the patient flow while testing various treatment guidelines. To estimate the Markov chains, we based our study on the EM algorithm presented by Yeh et al. (2010) to estimate the transition probabilities. The algorithm iteratively "completes" the sparse data and then estimates the Markov chain.

We simulate the steps of the treatment process for each patient, divided into two main time frames: events during an appointment with the physician and events between appointments. The guidelines suggest surveilling patients over 40 years old, prone to increase the chance of cardiovascular disease. The simulation ends if the patient has a CVD event, dies, or turns 80, assuming that physicians will immediately prescribe treatments. We test 4 different scenarios using a combination of cholesterol and blood pressure treatment guidelines. 1) (ACC 7.5%) Original ACC guidelines, 2) (ACC 5%) Using (5%) risk threshold for cholesterol, 3) (ACC 7.5%/JNC) Original ACC guideline for cholesterol and JNC for blood pressure, and 4) no treatment guidelines. We compare scenarios by estimating the percentage of patients with cholesterol-lowering treatment, percentage of patients with blood pressure-lowering treatment, and 10-year risk as a function of the patient's age.

3 RESULTS

We simulated patients who start at 40 years old with normal cholesterol and blood pressure levels and no treatment. We validated the simulation result by comparing it with the published estimate death rate by natural causes and the CVD event rate. We also performed a likelihood ratio test to evaluate the Markov chain assumption. After validation, we presented the results for male patients, as women were not sufficiently represented in our data set at all ages we considered. We divided male patients by race as is an important factor for 10-year risk prediction (Sussman et al. 2017).

We simulated each of the four scenarios referenced above and compared them. We notice that the ACC(5%) guideline starts treatment sooner, whereby at the age of 55, 14% more patients are on treatment than ACC(7.5%). Also, the JNC guideline tends to start medication much sooner. This result shows a much higher treatment burden if ACC(5%) and the JNC guidelines are used; however, there is no statistical difference in 10-year risk among the three scenarios. Therefore, we suggest using the original ACC guidelines for both cholesterol and blood pressure.

4 CONCLUSIONS

Health systems have improved the tracking of patients' medical data over time to the point where there is ample longitudinal data suitable for simulating health status changes over many years. Still, data may be missing or sparsely populated, which raises questions on how to build validated simulation models. This article presents a simulation model for patients treated to prevent CVD using a longitudinal data set from the U.S. VA health system. We tested different treatment guidelines for cholesterol and blood pressure to compare the risk reduction of heart attacks and strokes and treatment burden based on the proportion of patients in specific subgroups who were on medication at certain ages. We believe our study demonstrates a productive use of longitudinal EHR data that most health systems have collected. Moreover, the approach we took could be extended to other chronic diseases and conditions using EHR data. We were able to show that using EM-algorithms is a suitable approach to deal with sparse data in EHRs when building simulation models based on non-stationary Markov chains.

REFERENCES

Grundy, S. M., N. J. Stone, A. L. Bailey, C. Beam, K. K. Birtcher et al. 2018. "2018 AHA/ACC Guideline on the Management of Blood Cholesterol". *American College of Cardiology* 139(25):e1082–e1143.

James, P. A., S. Oparil, B. L. Carter, W. C. Cushman, C. Dennison-Himmelfarb et al. 2014, 02. "2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults (JNC 8)". *JAMA* 311(5):507–520.

Sussman, J. B., W. L. Wiitala, M. Zawistowski, T. P. Hofer, D. Bentley, and R. A. Hayward. 2017. "The Veterans Affairs Cardiac Risk Score". *Medical Care* 55(9):864–870.

Whelton, P., R. Carey, W. Aronow, D. Casey, and K. Collins. 2017. "2017 Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults". *American College of Cardiology* 15(3):127–248.

Yeh, H. W., W. Chan, E. Symanski, and B. R. Davis. 2010. "Estimating transition probabilities for ignorable intermittent missing data in a discrete-time markov chain". *Communications in Statistics: Simulation and Computation* 39(2):433–448.