Caring is Sharing – Exploiting the Value in Data for Health and Innovation
M. Hägglund et al. (Eds.)

doi:10.3233/SHTI230258

Modeling Patient Treatment Trajectories Using Markov Chains for Cost Analysis

Markus HAUG[a,1], Raivo KOLDE[a], Marek OJA[a] and Maarja PAJUSALU[a]
[a] *Institute of Computer Science, University of Tartu, Estonia*
ORCiD ID: Markus Haug https://orcid.org/0000-0003-0935-3307, Raivo Kolde https://orcid.org/0000-0003-2886-6298, Marek Oja https://orcid.org/0000-0002-3650-8194, Maarja Pajusalu https://orcid.org/0000-0001-9715-0678

Abstract. Electronically stored medical records offer a rich source of data for investigating treatment trajectories and identifying best practices in healthcare. These trajectories, which consist of medical interventions, give us a foundation to evaluate the economics of treatment patterns and model the treatment paths. The aim of this work is to introduce a technical solution for the aforementioned tasks. The developed tools use the open source Observational Health Data Sciences and Informatics Observational Medical Outcomes Partnership Common Data Model to construct treatment trajectories and implement these to compose Markov models for composing financial analysis between standard of care and alternatives.

Keywords. medical data, treatment trajectories, Markov chain, cost-effectiveness, OMOP CDM, heart failure

1. Introduction

The Observational Health Data Sciences and Informatics (OHDSI) Common Data Model (CDM) has been widely adopted for storing and analyzing medical observational data. It is currently being used in over 20 countries around the world. While the OHDSI CDM has a number of analytical tools available, there is a lack of tools specifically geared towards healthcare analytics with regards to costs and quality of life [1]. In response to this gap, we have developed easily distributable, transparent, and reproducibility-friendly research tools to fill this need.

2. Methods

To conduct healthcare analytics studies focused on costs and quality of life, we developed two R-packages. The first package, *Cohort2Trajectory*[2], utilizes the OHDSI CDM structure to generate patient trajectories by defining a target cohort and relevant

[1] Corresponding Author: Markus Haug, Research group of health informatics, https://health-informatics.cs.ut.ee; E-mail: markus.haug@ut.ee.
[2] https://github.com/HealthInformaticsUT/Cohort2Trajectory

state cohorts for the study. The tool resolves potential conflicts related to state overlap for each patient, offering various customization options for the decision-making process.

The second package, *TrajectoryMarkovAnalysis* [3], employs Markov modeling techniques on the resulting trajectory data to analyze the transition probabilities between health states. The package also retrieves and summarizes the costs associated with each trajectory, dividing them into state-specific costs. These Markov models and cost data are then used for both descriptive and cost-effectiveness analyses. The configurations for each study can be saved for reproducibility purposes.

3. Results and Discussion

To demonstrate the validity of our approach, we reproduced the study conducted by Thokala et al. [2], where they employed Markov models to calculate the incremental cost-effectiveness ratio for telemonitoring in heart failure patients. The results obtained using our developed tools on Estonian Health Insurance Fund data were similar to the original study. Specifically, the transition probabilities and monetary outputs of the Markov chains were comparable, leading to the same non-supporting conclusion about the feasibility of using telemonitoring for heart failure patients due to the high incremental cost. The calculated cost per quality-adjusted life year (QALY) exceeded the cost-effectiveness threshold agreed upon by healthcare institutions, making telemonitoring not a viable option for heart failure patients. Additionally, we employed survival analysis to compare the observed and generated trajectories, validating the learned Markov models.

The methods used to develop our tools have demonstrated their effectiveness and utility. By distributing our packages among OHDSI data partners, we can facilitate the reproducibility and transparency of any study defined using these tools. One potential challenge in conducting such studies is the availability and reliability of monetary data among data partners. Moving forward, we plan to focus on improving validation and comparison methods, including addressing potential issues related to the use of monetary data. Additionally, we aim to explore other methods for comparing patient treatment trajectories, including clustering trajectories with dynamic time warping and predicting future trajectories with neural networks.

The ability to analyze observational trajectories has numerous potential applications, including evaluating drug adherence and monitoring compliance or changes in the standard of care. In conclusion, our research emphasizes the importance of distributable and reproducible tools in healthcare for advancing the field.

References

[1] Petitti Kent S, Burn E, Dawoud D et al. Common Data Model Solutions: Evidence Generation for Health Technology Assessment. PharmacoEconomics. 2021; 39:275–285, doi: https://doi.org/10.1007/s40273-020-00981-9

[2] Thokala P, Dodd P, Baalbaki H, Brennan A, Dixon S, Lowrie K. Developing Markov Models From Real-World Data: A Case Study of Heart Failure Modeling Using Administrative Data. Value in Health. 2020; 23(6):743–750, doi: https://doi.org/10.1016/j.jval.2020.02.012

---

[3] https://github.com/HealthInformaticsUT/TrajectoryMarkovAnalysis