PCSI 2025
Québec, Canada
September 9 to 12

Developing and Improving Predictive Models of the Risk of Health Utilisation and Future Healthcare expenditures: National Recalibrations and new markers

Authors: Stephen Sutch (a), Klaus Lemke (a), Harriet Martyn (b)

Introduction

Several models predict the risk of hospitalization and healthcare expenditures from general and insured populations. These models are used for various purposes, including screening patients for Case Management Programs, Disease Management Programs, organizational profiling, and assessing financial risk. The importance of locally calibrated models has been previously discussed, along with the use of primary care and community data. Social determinants of health (SDoH) and behavioural and routine health measurement data are also crucial but often underutilized in predictive models.

Methods

The predictive models were derived using patient-level data, with classification of diagnostic, pharmaceutical, and historic utilization data, using the Johns Hopkins ACG System to reduce variables and provide measures of multimorbidity. Logistic and Linear Regressions were used to produce models on outcomes such as hospitalization within 12/6 months, emergency/unplanned hospitalization within 12 months, and healthcare expenditures in the preceding 12 months. Alternative modelling methods, including non-negative models and Lasso Regression, Random Forests, Gradient Boosting Models were also explored. Validation of models is essential to test accuracy, efficacy, and bias.

Results

Results will be shown from multiple general populations. Intermediate classifications were updated in the process of redeployment of models. Both changes in existing classifications and the addition of new variables, such as Social Need, Patient Need and behavioural. Although the original models generalise well to other populations, the recalibrated models using local data produce better performance.

Discussion

Comprehensive person-based records are important for such models, especially with health policy oriented towards integrated care. Local recalibration ensures models are relevant to the population they will be applied to and provide better performance than the original models. Traditional modelling techniques like logistic and linear regression can efficiently create these models and provide good face validity for local users, but alternative models are being explored. Casemix classifications reduce data complexity and provide robust measures of key constructs like multimorbidity.

There is increased interest in recognizing earlier and emerging risk for more preventative methods, such as chronic disease self-management programs. Emerging data from Electronic Health Records (EHR), Personal Health Records (PHR), and Social Care and behavioural data are expected to provide greater insight into populations with the highest need.

References

01 Corti MC, Avossa F, Schievano E, et al. A case-mix classification system for explaining healthcare costs using administrative data in Italy. Eur J Intern Med 2018;54:13-6.:S0953-6205(18)30095-5. doi:10.1016/j.ejim.2018.02.03. https://www.sciencedirect.com/science/article/abs/pii/S0953620518300955 02 Zachary Predmore, Elham Hatef, Jonathan P. Weiner. Integrating Social and Behavioral Determinants of Health into Population Health Analytics: A Conceptual Framework and Suggested Road Map. Population Health Management Vol. 22, No. 6 Dec 2019. https://www.liebertpub.com/doi/abs/10.1089/pop.2018.0151

03 Girwar SM, Fiocco M, Sutch SP, Numans ME, Bruijnzeels MA. Validating and Improving Adjusted Clinical Group's Future Hospitalization and High-Cost Prediction Models for Dutch Primary Care. Popul Health Manag. 2023 Dec;26(6):430-437. doi: 10.1089/pop.2023.0162. Epub 2023 Nov 2. PMID: 37917048. https://www.liebertpub.com/doi/pdf/10.1089/pop.2023.0162

a: Johns Hopkins University, United States
b: Johns Hopkins HealthCare Solutions, United States