



High-Risk Patient Identification: A Review of Current Methodologies

Shalini Sivanandam¹ and Bradley N. Doebbeling^{1,2}

¹Program of Biomedical Informatics, Arizona State University, USA

²Program of Science of Health Care Delivery, Arizona State University, USA

*Corresponding author: Shalini Sivanandam, Program of Biomedical Informatics, College of Health Solutions, Arizona State University, Phoenix, AZ, USA

To Cite This Article: Shalini Sivanandam. High-Risk Patient Identification: A Review of Current Methodologies. *Am J Biomed Sci & Res.* 2019 - 6(1). *AJBSR.MS.ID.000999*. DOI: [10.34297/AJBSR.2019.06.000999](https://doi.org/10.34297/AJBSR.2019.06.000999).

Received: 📅 September 26, 2019; Published: 📅 November 06, 2019

Abstract

Identifying High Cost High Need (HCHN) patients has been increasingly identified as important among healthcare stakeholders since the 1997 Balanced Budget Act prompted the Health Care Financing Administration (HCFA) to require risk-adjusted payment methodologies. High risk patient identification was originally derived primarily from claims data. This has since evolved to using multiple data sources, including clinical data from electronic health records, self-reported health measures etc.

Importantly, the case definition for HCHN patients is evolving and needs further research. Various high risk identifications are used by payers and providers for a range of purposes ranging from targeted outreach for disease management, reducing readmissions, and strategies for greater care coordination through patient centered medical homes and Accountable care organizations (ACOs). However, the great diversity of case definitions, diversity among stakeholders, availability of source data, access to technology and analytical manpower, all complicate the refinement and use of a high-risk patient identification.

With the ongoing momentum to leverage big data technologies, several solutions for HCHN patient identification are being developed. It is therefore imperative to understand the current practices and evidence from research prior to adopting a flexible and proactive predictive analytics approach to HCHN patients. The critical need for further research and funding is outlined.

Keywords: High risk; High Cost High Need; Predictive modelling; Risk adjustment; Research priorities

Abbreviations: HCHN: High Cost High Need; VBP: Value Based Payment; ACO: Accountable Care Organizations; HER: Electronic Health Records

Introduction

Healthcare has been undergoing a transition to meet the Triple Aim goals of: improved patient care experience, population health and reduced cost through value-based payment (VBP) incentives. Payment models that risk adjust for patient case mix have been used increasingly for reimbursement of healthcare providers. Incentives from value-based contracting encourage collaborative care models among providers in various care settings to assume responsibility for vulnerable patient populations [1]. To reduce costs and improve health outcomes, these efforts create care pathways for high-risk patients that can require active monitoring and care planning.

Such initiatives place the burden of high-risk member identification on the providers, while payers still employ rudimentary risk adjusted payment models. Collaborative care

models sometimes provide services not reimbursable by payers even though they may yield long term benefits [2,3]. In response, many organizations are developing care models for chronic disease states to risk stratify their patient population and employ a targeted approach at care coordination to conserve costs [3,4]. Since traditional CMS risk adjustment models over rely on historical data with significant limitations, the industry is responding with meaningful ways to incorporate big data and sophisticated machine learning algorithms with mixed results [3].

Materials and Methods

Articles were identified using search terms: High risk patient identification, predicting high cost patients, high cost high need patient identification, risk adjustment analysis, predictive analytics,

risk assessment, risk adjustment for healthcare payment. The search was limited to peer-reviewed articles since the year 2000, to maintain relevance to current day practices. Results were then screened on title and abstract, with a focus primarily on predictive modeling methodology and secondarily on background of risk adjustment policy and practice. The resulting papers were then reviewed in detail with full text assessment to evaluate their potential impact. Here we summarize the key findings from the highest impact articles.

Results and Discussion

Defining High risk patients

According to an expert panel from the Commonwealth Fund, high-need and high-cost (HCHN) patients are typically those with complex clinical conditions that limit their ability to care for themselves and should be the focus of federal initiatives to reduce costs [5,6]. The top 5% of most costly patients contribute to approximately 50% of healthcare costs as shown consistently in literature. High-cost and high-need patients include patients with three or more chronic diseases with functional limitations that impact their self-care and routine activities of daily living. Most definitions include a behavioral or psychological component as one of the chronic diseases, and social determinants of health as well [7].

Main aspects of predictive analytics

Some investigators consider adverse health events are random occurrences and question the validity of using previous utilization and cost data to predict future outcomes. It is theorized that those who have a serious medical event and receive active intervention, are typically are not the same that contribute to costs in the near future [8]. However, most would agree that individuals with multiple medical and behavioral comorbid chronic conditions and unmet social needs are at increased risk for adverse health events such as emergency department visits, hospitalizations and surgery. Models that rely on identifying patients in the upper quartiles of cost or utilization don't account for the principles of regression to the mean, i.e. most high-cost patients will improve even without intervention [8]. In contrast, predictive analytics that address the composite impact of various patient, organizational or system factors that predict adverse events and utilization can have superior results.

Claims vs. clinical data: A significant limitation of healthcare data is that it is largely incomplete or inaccurate. Claims data is primarily generated for administrative and reimbursement purposes, whereas clinical documentation is an inexact source of data with significant user variability [8]. The literature shows there are evident limitations to relying on claims data in high-risk prediction, since administrative data is not representative of clinical risk [9]. Contrastingly, even though there are significant

advantages to using EHR data, and review of literature shows that they are underutilized [10].

Despite the limited availability of clinical data, there are incentives for providers to leverage EHR technology to make this data available. However, claims data is readily available and far more standardized than clinical data and most contemporary risk models are dependent on claims data. Comparing risk assessment tools that used a. patient demographics, b. self-reported outcomes and claims based methods to predict costs, claims based methods to be still more effective than the alternatives [11]. EHR data, even when available is frequently unstandardized and user dependent, and often incomplete or inaccurate [8].

Big data methodologies: Although machine learning algorithms are increasingly used in healthcare, barriers to wider application include: 1. difficulties with feature selection, especially with respect to determining temporal correlation and 2. Inadequate explanation of prediction results that limits actionable clinical applications. The slow adoption of machine learning is also attributed to the limited expertise of analysts in applying and interpreting these complex methodologies [12]. Kan et al., found making incremental changes to standard linear regression models, with minor modifications could yield better results when knowledge and resources are constrained. While drawing causal inference and unbiased estimates from machine learning models may be ineffective, payers and providers can still leverage them to create robust prediction models to identify high-risk.

Key limitation: Generalizability: While various predictive models have been developed to identify the HCHN patient, the generalizability of the methods is limited by the type of care setting and the healthcare specialty area that delivered the patient care studied. Especially, when administrative claims data is the main data source, with variations in socioeconomic, clinical, lab and medication data collected in EHR, the results of most predictive methodologies are not applicable across care settings and patient populations. The main limitation of several existing models are they are skewed to the target populations and have marginal generalizability.

Conclusions

Further research is needed on a wide range of issues, including the refinement of common definitions of the HCHN patient and identifying strategies to improve their care that are generalizable. The potential opportunity for policy makers and providers in developing a scalable predictive model for a high yield set of chronic conditions and social determinants of health is of significant value. Much of the current research on predictive modeling offers limited information on the machine learning methodology and are challenging to evaluate [13,14]. Frameworks that offer proof of concepts or standard approaches with explicit open source code and best practices need to be made available. Investments

in advancing the techniques for predictive modeling are clearly critical. Further research that addresses the constantly evolving needs of the industry are imperative.

Acknowledgements

We thank Dr. William Johnson for his help with formalizing the research topic.

Conflict of Interest

The authors note no conflicts of interest.

References

1. Tianzhong Y, Yang Y, Yugang J, Xiao L (2019) Dynamic prediction of hospital admission with medical claim data. *BMC Medical Informatics and Decision Making* 19(S1): 1-14.
2. Bates DW (2015) Health Information Technology and Care Coordination: The Next Big Opportunity for Informatics? *Yearbook of medical informatics* 10(1): 11-14.
3. Gillespie G (2015) HOW ANALYTICS TAKES THE UNKNOWN OUT OF ACOs. *Technology helping organizations connect the dots and provide better care. Health data management* 23(4): 18-22.
4. Wammes JJG, van Der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT (2018) Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ open* 8(9): e023113.
5. Blumenthal D, Abrams MK (2016) Tailoring Complex Care Management for High-Need, High-Cost Patients. *JAMA* 316(16): 1657-1658.
6. Hayes SL, Salzberg CA, McCarthy D, Radley DC, Abrams MK, et al. (2016) High-Need, High-Cost Patients: Who Are They and How Do They Use Health Care? A Population-Based Comparison of Demographics, Health Care Use, and Expenditures. *Issue brief (Commonwealth Fund)* 26: 1-14.
7. Long P, Abrams M, Milstein A, G Anderson, K Lewis Apton, et al. (2017) Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health. Washington, DC: National Academy of Medicine.
8. Cousins MS, Shickle LM, Bander JA (2004) An Introduction to Predictive Modeling for Disease Management Risk Stratification. *Disease Management* 5(3): 157-167.
9. Yang C, Delcher C, Shenkman E, Ranka S (2018) Machine learning approaches for predicting high cost high need patient expenditures in health care. *Biomedical engineering online* 17(Suppl 1): 131.
10. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA (2017) Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *Journal of the American Medical Informatics Association: JAMIA* 24(1):198-208.
11. FitzHenry F, Shultz EK (2000) Health-risk-assessment tools used to predict costs in defined populations. *Journal of healthcare information management: JHIM* 14(2): 31-57.
12. Kan HJ, Kharrazi H, Chang HY, Bodycombe D, Lemke K, et al. (2019) Exploring the use of machine learning for risk adjustment: A comparison of standard and penalized linear regression models in predicting health care costs in older adults. *PLoS One* 14(3): e0213258.
13. Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR (2018) Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *Journal of the American Medical Informatics Association: JAMIA* 25(8): 969-975.
14. Rose S (2016) A Machine Learning Framework for Plan Payment Risk Adjustment. *Health Services Research* 51(6): 2358-2374.