Article published online: 2024-01-22

Perspective

Enhancing Hospitalized Patients' Palliative Care Referrals via Machine Learning-Based Predictive Modeling within Electronic Health Record Systems

Arun Ghoshal[1]

[1] Princess Margaret Cancer Centre, Toronto, Canada

Ind J Med Paediatr Oncol

Address for correspondence Arun Ghoshal, MD, MRes, Clinical Research Fellow, Princess Margaret Cancer Centre, 610 University Avenue, Toronto ON M5G 2M9, Toronto, Canada (e-mail: arunghoshal@outlook.com).

Abstract

Access to palliative care (PC) holds significance for hospital-admitted patients grappling with the symptoms of life limiting illnesses. Nonetheless, numerous such patients who could gain from PC fail to receive it promptly or even at all.

We can leverage the prior year's historical data extracted from electronic health records of hospitalized patients to train a machine learning (ML) model. This model's purpose would be to prognosticate the requirement for PC consultation using real-time data. The model, operating as a semi-supervised system, will be integrated into institutional data pipelines, and utilized by a downstream display application overseen by the PC team. In cases where the PC team deems it suitable, a team member will communicate with the respective care team of the patient. The ML model's training efficacy will be assessed using the area under the curve (AUC) metric, employing a 20% reserved validation set. The threshold for PC consultations will be grounded in historical data. To enhance the ML model's precision, the pivotal variables within the model will be pinpointed, and any sources of biases or errors in the model will be identified for meticulous refinement. The AUC values of successive ML models will be juxtaposed with cross-validation data.

Automatizing the referral procedure through electronic health record systems has the potential to usher in a more effective and streamlined approach to healthcare delivery.

Keywords

- hospitalized patients
- palliative care referrals
- machine learning
- predictive modeling
- electronic health records

Introduction

In recent years, there has been a growing recognition of the importance of palliative care (PC) in improving the quality of life for hospitalized patients with serious illnesses. PC focuses on providing relief from symptoms, pain, and stress, aiming to improve the overall well-being of patients and their families. However, despite its significant benefits, there is a notable gap in the timely identification and referral of patients who could benefit from PC services.[1]

To address this challenge, there is an emerging opportunity to leverage the power of machine learning (ML)-based predictive modelling within electronic health record (EHR) systems.[2] By harnessing the vast amount of patient data stored in EHRs, it is possible to develop predictive models that can identify hospitalized patients who are most likely to benefit from PC interventions.

This concept article aims to explore the potential of using ML techniques to enhance the referral process for PC within hospital settings. By integrating predictive models into EHRs,

DOI https://doi.org/10.1055/s-0043-1776357.
ISSN 0971-5851.

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

healthcare providers can identify patients who may benefit from early PC interventions, allowing for more proactive and personalized care planning. In contrast to prior methods, this strategy aims to directly forecast PC consultations instead of relying on mortality as an indirect measure. The outcomes of the model can subsequently be incorporated into the procedures of the PC team.

This article will first provide an overview of the current challenges in PC referrals and the potential benefits of leveraging ML algorithms in addressing these challenges. It will then discuss the key components and considerations involved in developing a ML-based predictive model for PC referrals, including data collection, feature selection, model training, and evaluation. Furthermore, ethical and privacy considerations will be examined, as the use of patient data in predictive modelling raises important concerns regarding confidentiality and informed consent. This article will propose strategies to mitigate these concerns and ensure the responsible and ethical use of patient data. Finally, this concept article will highlight potential implementation challenges and opportunities, as well as the expected impact of integrating ML-based predictive modelling into EHRs for PC referrals.

Overview of the Current Challenges in PC Referrals and the Potential Benefits of Leveraging ML Algorithms

PC referrals face several challenges within hospital settings, resulting in delayed or missed opportunities for patients to receive the appropriate care. Some of the key challenges include subjective patient identification reliant on healthcare providers' assessments, delayed referrals, lack of standardized referral criteria, and resource constraints.[3] Leveraging ML algorithms within EHRs offer promising opportunities to address the challenges in PC referral triggers, which could include diagnosis of life-limiting or terminal illnesses (e.g., advanced cancer, end-stage heart failure), severe and progressive symptoms (e.g., uncontrolled pain, dyspnea), frequent hospitalizations and readmissions, decline in functional status and quality of life, and limited response to curative treatments.[4–8] Some potential benefits include early identification and intervention by analyzing patient data available in EHRs, objective and standardized referral criteria, enhanced accuracy and efficiency, and personalized care planning as ML models can consider a wide range of patient characteristics, such as medical history, comorbidities, and PC needs assessment like physical assessment of symptoms (pain, nausea, fatigue), psychosocial assessment (emotional distress, spiritual concerns), discussion of patient's goals, values, and preferences, and assessment of family and caregiver needs to develop personalized care plans.[2,9]

Clinical Use Case Design

The intended model aims to predict consultations with PC for patients during their hospitalization. This prediction will facilitate prompt PC availability, potentially enhancing the satisfaction of both patients and caregivers with the healthcare provided. The model will operate as a web-based service, integrated with institutional data flows, and utilized by a subsequent display application overseen by the PC team. For those patients that the PC team deems appropriate, a team member will then contact the patient's corresponding care team. Stakeholders will be engaged when designing the solution (both for buy-in and support and for end-user design) and will include hospital administration, non-PC healthcare teams, medical informatics, the information technology team, patients/caregivers, and funders. Outcomes of enhanced PC referrals would include improved symptom management and pain control, enhanced patient and family satisfaction, better alignment of medical care with patient values and preferences, reduced hospital readmissions and emergency department visits, and enhanced communication and coordination among healthcare providers.[9] Often these might lead to cost savings through reduced hospitalizations, decreased utilization of expensive interventions, improved resource allocation and healthcare efficiency, and enhanced patient and family satisfaction leading to better reputation for the hospital.[10,11]

Design Thinking Methodology

We will start with involve the users to understand the needs and challenges of the stakeholders involved, including patients, caregivers, healthcare providers, and PC specialists via interviews, surveys, and observations to gain insight into their experiences and pain points. We will ideate solutions via brainstorming over a wide range of workable solutions to the problem, without worrying about feasibility at this stage, via mind mapping, brainstorming sessions, and SCAMPER (Substitute, Combine, Adapt, Modify, put to other uses, Eliminate, Rearrange) to generate ideas,[12] and diverse perspectives and challenges to assumptions will be encouraged at this stage. Next step involves developing a rough prototype of a web application, using wireframes, sketches, or low-fidelity mock-ups. We will perform user testing to get feedback on prototypes using techniques like A/B testing or surveys to get quantitative data on user preferences.[13] We will institute necessary changes based on the feedback to the web application and repeat the prototyping and testing process until we have a final product that meets the needs of users and achieves goals.

Data and Pre-Processing

We will get the necessary permission to collect hospital data for 1 year, for adults (>18 years), and inpatients, excluding patients in the PC unit or waiting to be transferred to an external hospice. The data category will encompass numeric, categorical, or binary formats contingent upon the nature of the variable. Instances consist of patient demographic information, past utilization (quantified by tallying recent instances of inpatient care, intensive care, and primary care interactions before admission), concurrent health conditions (ICD11 diagnosis codes), and dynamic data like lab results and the ongoing length of stay.

Data Ingestion and Preprocessing

In the existing framework of system architecture, whenever a modification is applied to EHRs, the system initiates the creation of HL7 messages. These messages are then disseminated within the organization through the Enterprise Service Bus(ESB) version.[14] The ESB is subject to ongoing surveillance by the Project Data Pipeline. This pipeline includes a rules manager designed to implement clinical enrichment rules onto the messages. The outcomes produced by this rules manager can encompass significant data required for the operation and presentation of the application. Alternatively, if necessary, it can also trigger predictive processing in situations where changes in pertinent variables for monitored patients come into play.

Model Development and Validation

We intend to create a semi-supervised learning framework by utilizing both labeled datasets and unlabeled outcomes.[15] Specifically, our focus is on developing a model that employs a Poisson distribution process to mathematically capture the concept of time-to-event outcomes, particularly in the context of PC consultations.[16] These consultations can be understood as events occurring at varying rates over time. Our approach involves utilizing a classification task with a decision tree algorithm.

To construct the response variable, we will measure the time from admission to a PC consult, taking into account that this measurement is right-censored at the discharge time. The Poisson process will enable us to predict the rate of PC interventions per unit of time. For practical implementation in a clinical setting, we will transform this rate into a 7-day probability of PC consultation. While various ML models can be employed that can leverage Poisson likelihood, our choice is the Gradient Boosting Machine (GBM). The GBM's ability to incorporate a loss function equivalent to the minus log-likelihood for Poisson distributions makes it well-suited for our purpose.

Furthermore, the GBM can effectively handle missing values, which is particularly important considering the presence of missing data in laboratory values and diagnosis codes. Our model will be dynamic, updating predictions as new information becomes available through a carry-it-forward dataset construction approach. However, it is important to acknowledge that assuming constant predictor values between observations might introduce bias or errors.

Upon exceeding a predefined threshold, the model will trigger notifications to the PC team. The threshold will categorize the 7-day probability of PC consultation as "low," "medium," or "high." The "high" category will be calibrated based on the existing capacity of the PC service, aiming to align with the average of around ten new consults daily, including those via the traditional pathway. This calibration will involve utilizing data to establish a threshold for the receiver operating characteristic (ROC) curve.

The labeled dataset will be divided into 80% for training and 20% for testing. We will determine hyperparameters for the GBM model—such as the number of trees, shrinkage, and interaction depth—through cross-validation. This process entails partitioning the dataset into random subsets, using one as the test set while training on the others. Repeating this process for each subset, the average results will guide our final choice of hyperparameters. We will employ 5- to 10-folds, depending on our computational resources. The optimal hyperparameters will be those producing the highest area under the curve (AUC) in the time-dependent ROC curve. This curve will plot AUC and positive predictive value.

Ultimately, the final model will be fitted using all available data and evaluated in real-time using unlabeled data. Its performance will be compared to the cross-validation data, allowing us to assess the AUC and make any necessary adjustments to the model.

User Validation and Clinical Integration

Validation metrics will include the number of PC consults triggered by the ML model, compared with previous years' data, and changes in clinical outcomes guided by the new ML model. End users will be internal stakeholders, to begin with, or external customers if successfully implemented. We will gather feedback from users within the time frame for testing and analyze patterns in the feedback and identify areas where the model can be improved. We will make improvements like adjusting the model parameters, collecting additional data, or making changes to the user interface. We will repeat the process until satisfied with the performance of the ML model.

We will identify key stakeholders who will be impacted by the introduction of the ML model, such as clinicians, administrators, patients, and IT staff. Engage with them early in the process to understand their needs and concerns and involve them in the decision-making process. We will clearly articulate the benefits of the ML model to stakeholders like increasing efficiency, reducing costs, and enhancing the quality of care. We will anticipate and address concerns and potential risks associated with the introduction of the ML model. Communicate clearly about the safeguards and measures in place to protect patient privacy and data security. Provide evidence-based information to address concerns about the accuracy and reliability of the model. We will provide adequate training and support to stakeholders to ensure they are comfortable with the technology and understand how to use it effectively. Offer ongoing support to address any issues or questions that may arise. We will monitor and evaluate the implementation of the ML model to ensure it is meeting the desired outcomes and identify areas for improvement. Share the results with stakeholders to demonstrate the impact of technology and foster ongoing support and engagement.

Ethics, Legal, and Regulatory Considerations

Transparency: It is important to ensure that the ML application is transparent in how it makes its decisions. Patients and healthcare professionals need to be able to understand how the algorithm arrived at its recommendations, so they can trust and act upon them. This holds significant significance in the context of end-of-life choices. To illustrate, comprehending and explaining the course that a decision tree follows to arrive at its conclusion is straightforward, whereas tracing the trajectories of numerous trees, numbering in the hundreds or thousands, becomes notably more challenging.

Autonomy: The ML application should not replace the judgment of healthcare professionals. It should be designed to support clinical decision-making, rather than replace it. It is up to the healthcare professional to make decisions about the care of their patients.

Accountability: It is important to ensure that the ML application is accountable for its decisions. This means that there should be mechanisms in place to monitor its performance and identify any errors or biases that may arise.

Informed consent: Patients should be fully informed about the use of ML in their care and should have the option to opt out if they choose. This requires clear communication about how the ML application works and the implications of its recommendations.

Privacy: The use of personal data is essential for ML algorithms to work, which can potentially reveal sensitive medical information about patients. Therefore, it is important to ensure that patient privacy is protected throughout the development, deployment, and use of the ML application.

Discussion and Conclusion

This concept article aimed to explore a project utilizing the potential of ML-based predictive modelling within EHRs to enhance the referral process for PC in hospitalized patients, allowing for timely and appropriate referrals. The accuracy and performance metrics achieved by these models will suggest if they have the potential to significantly improve the current referral process. Proactive identification of patients who may benefit from PC can initiate early conversations and interventions, leading to better patient outcomes and enhanced quality of life.

One of the key advantages of using ML in this context would be the ability to leverage a wide range of patient data, including demographics, clinical variables, laboratory results, and diagnostic codes. These models can detect patterns and associations that may not be apparent to human clinicians, enabling more accurate predictions. Furthermore, the models can continuously learn and adapt from new data, refining their predictions over time and improving their performance.

However, several challenges and limitations need to be addressed before implementing ML-based predictive modelling for PC referrals in electronic health systems.

Primarily, ensuring data quality and standardization across different healthcare settings is crucial. Discrepancies, absent information, and mistakes within EHRs can result in prejudiced or erroneous forecasts, posing a potential risk to patient well-being. To illustrate, the training of the model will utilize a historical cohort of primary care consultations that is not flawless; specifically, numerous individuals who could have gained from a primary care consultation might not have undergone one. This disparity has the potential to cause oversights by the algorithm, especially if there exists a consistent bias towards primary care consultations under the existing protocols. Mitigation strategies may include effort to get better representative data, using metrics like precision, recall, F1 score, adjust class weights, fairness constraints in model training, regular model audit, and monitoring. Therefore, data preprocessing and quality assurance processes must be robustly implemented.

An additional obstacle pertains to the comprehensibility of ML models. Despite their ability to attain remarkable predictive precision, these models frequently lack transparency in elucidating the fundamental decision-making process. This absence of clarity could impede the establishment of trust and approval from healthcare experts. To illustrate, consider the presumption of carrying forward data for time-dependent predictors, where the assumption that a patient's hemoglobin level remains constant for 3 days solely because a laboratory test was conducted on a Monday and not repeated until Thursday does not hold true. The advancement of interpretable ML techniques, capable of offering insights into the drivers behind predictions, is a domain that demands further attention.

Additionally, ethical considerations must be carefully addressed when implementing ML models in healthcare. Patient privacy, data security, and informed consent are of utmost importance. Safeguards must be in place to protect sensitive patient information and ensure compliance with privacy regulations. It is crucial to strike a balance between the potential benefits of predictive modelling and the ethical responsibilities associated with its use. Should our solution become part of a clinical workflow, obtaining approval from the Food and Drug Administration (FDA) could be necessary. For instance, the US FDA categorizes medical devices, including software, into four different classes, depending on the level of risk they pose to both the patient and the user.[17]

Despite these challenges, the potential impact of ML-based predictive modeling in enhancing PC referrals within EHRs is significant. Future research should focus on prospective validation of the models in real-world clinical settings to assess their performance in diverse patient populations. Additionally, studies should explore the impact of implementing these models on clinical workflow, patient outcomes, and healthcare resource utilization.

Furthermore, collaboration between clinicians, data scientists, and policymakers is essential for successful implementation. Engaging healthcare providers in the development and validation of these models can help address concerns, foster trust, and ensure that the technology aligns with clinical needs and workflows. Moreover, policymakers should establish guidelines and regulations that promote the responsible and ethical use of ML in healthcare.

In conclusion, the use of ML-driven predictive modeling shows significant potential in improving the process of referring patients for PC during their hospitalization. Detecting individuals who could gain from PC at an earlier point can result in better patient results and more effective utilization of resources. Nevertheless, it is crucial to meticulously address concerns related to data accuracy, comprehensibility, and ethical implications. By tackling these obstacles and delving deeper into the possible advantages, we can make strides in incorporating ML into EHRs, thereby enhancing the provision of PC for those requiring it most.

Patient Consent

None Declared.

Disclosure of Funding Support

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The author(s) declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Artificial intelligence for Clinician Champions Certificate Program, The Michener Institute of Education at UHN.

References

1 Palliative care. Accessed October 8, 2023 at: https://www.who.int/news-room/fact-sheets/detail/palliative-care
2 Ghosheh GO, Thwaites CL, Zhu T, Thwaites CL, Zhu T. Synthesizing electronic health records for predictive models in low-middle-income countries (LMICs). Biomedicines 2023;11(06):1749
3 Barry C, Paes P, Noble S, Davies A. Challenges to delivering evidence-based palliative medicine. Clin Med (Lond) 2023;23(02):182–184
4 Preisler M, Rohrmoser A, Bar K, Letsch A, Goerling U. Early integration of palliative/supportive cancer care-healthcare professionals' perspectives on the support needs of cancer patients and their caregivers across the cancer treatment trajectory. Oncol Res Treat 2017;40(Suppl 3):187
5 Grant M, Elk R, Ferrell B, Morrison RS, von Gunten CF. Current status of palliative care–clinical implementation, education, and research. CA Cancer J Clin 2009;59(05):327–335
6 Kristjanson LJ, Aoun SM, Oldham L. Palliative care and support for people with neurodegenerative conditions and their carers. Int J Palliat Nurs 2006;12(08):368–377
7 Radbruch L, De Lima L, Knaul F, et al. Redefining palliative care-a new consensus-based definition. J Pain Symptom Manage 2020;60(04):754–764
8 Ramasamy Venkatasalu M, Sirala Jagadeesh N, Elavally S, Pappas Y, Mhlanga F, Pallipalayam Varatharajan R. Public, patient and carers' views on palliative and end-of-life care in India. Int Nurs Rev 2018;65(02):292–301
9 Stiel S, Pastrana T, Balzer C, Elsner F, Ostgathe C, Radbruch L. Outcome assessment instruments in palliative and hospice care–a review of the literature. Support Care Cancer 2012;20(11):2879–2893
10 May P, Garrido MM, Cassel JB, et al. Cost analysis of a prospective multi-site cohort study of palliative care consultation teams for adults with advanced cancer: where do cost-savings come from? Palliat Med 2017;31(04):378–386
11 Morrison RS, Penrod JD, Cassel JB, et al; Palliative Care Leadership Centers' Outcomes Group. Cost savings associated with US hospital palliative care consultation programs. Arch Intern Med 2008;168(16):1783–1790
12 Michalko M Thinkertoys: a handbook of creative-thinking techniques.:355. Accessed October 8, 2023 at: https://books.google.com/books/about/Thinkertoys.html?id=y7O_6v1c52wC
13 Kohavi R, Longbotham R Online Controlled Experiments and A/B Tests. Encyclopedia of Machine Learning and Data Science. Published online 2023:1–13. Doi: 10.1007/978-1-4899-7502-7_891-2
14 Health Information Systems: Concepts, Methodologies, Tools, and Applications ... - Google Books. Accessed October 8, 2023 at: https://books.google.ca/books?id=WnBJsRtfVbYC&pg=PR39&redir_esc=y#v=onepage&q&f=false
15 Novak PK, Lavrač N, Webb GI. Supervised descriptive rule induction. Encyclopedia of Machine Learning and Data Mining 2016:1–4
16 Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013;7(DEC):21
17 Software as a Medical Device (SaMD) | FDA. Accessed October 8, 2023 at: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd

