ChiKD CHILDHOOD KIDNEY DISEASES

Special Article

Child Kidney Dis 2024;28(3):93-98
pISSN 2384-0242 · eISSN 2384-0250
https://doi.org/10.3339/ckd.24.018

Integrating predictive modeling and causal inference for advancing medical science

Tae Ryom Oh[1]

[1]Department of Internal Medicine, Mokpo Hankook Hospital, Mokpo, Republic of Korea

Artificial intelligence (AI) is revolutionizing healthcare by providing tools for disease prediction, diagnosis, and patient management. This review focuses on two key AI methodologies in healthcare: predictive modeling and causal inference. Predictive models excel in identifying patterns to forecast outcomes but are limited in explaining the underlying causes. In contrast, causal inference focuses on understanding cause-and-effect relationships, which makes effective medical interventions possible. Although randomized controlled trials (RCTs) are the gold standard for causal inference, they face limitations including cost and ethical concerns. As alternatives, emulated RCTs and advanced machine learning techniques have emerged for estimating causal effects, bridging the gap between prediction and causality. Additionally, Shapley values and Local Interpretable Model-Agnostic Explanations improve the interpretability of complex AI models, making them more actionable in clinical settings. Integrating prediction and causal inference holds great promise for advancing personalized medicine, enhancing patient outcomes, and optimizing healthcare delivery. However, careful application of AI tools is crucial to avoid misinterpretation and maximize their potential.

Keywords: Artificial intelligence; Causality; Forecasting

Introduction

Artificial intelligence (AI) is transforming healthcare and also advancing early disease detection, personalized treatment, and operational efficiency [1-6]. To fully leverage the potential of AI, it is crucial to differentiate between two primary categories of healthcare applications, prediction and causal inference, each requiring different methodologies.

Currently, AI is primarily used for prediction tasks [7]. Predictive models forecast future outcomes based on historical data, and identify patterns and correlations. These models are valuable for predicting patient readmission and chronic disease progression [8,9]. However, they have limitations, particularly in healthcare, where understanding the root cause of a condition is vital for effective treatment. As the role of AI in healthcare grows, the importance of causal inference is increasingly being recognized [10]. Unlike prediction, which focuses on what may happen, causal inference seeks to determine why something happens by identifying cause-and-effect relationships. This understanding is key for developing targeted interventions [11-13]. Although the predictions and causal inferences are complementary, they are not interchangeable. Both offer unique advantages in specific contexts (Fig. 1). The confusion of predictions with causal inferences can lead to significant errors and thus compromise patient care [14].

By appropriately applying both prediction and causal in-

Received: October 1, 2024; Revised: October 23, 2024; Accepted: October 24, 2024

Correspondence to
Tae Ryom Oh
Department of Internal Medicine, Mokpo Hankook Hospital, 483 Yeongsan-ro, Mokpo 58643, Republic of Korea
E-mail: tryeom.oh@outlook.com

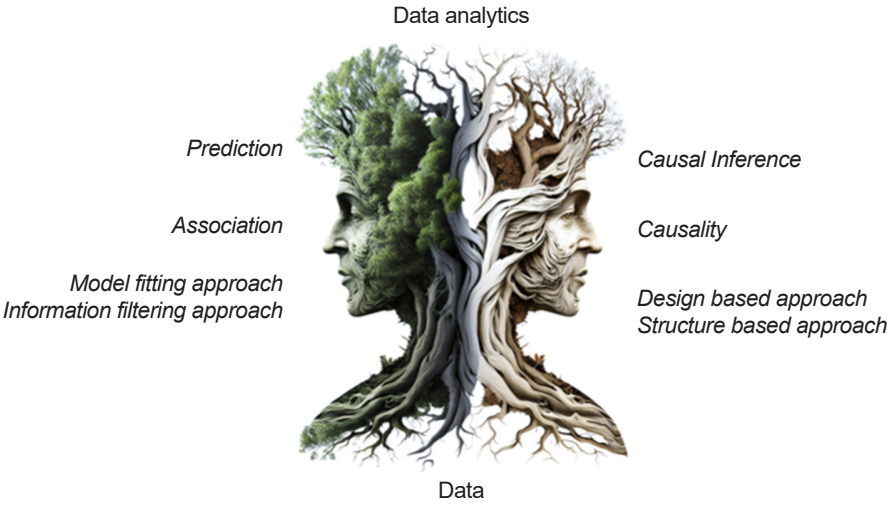

Fig. 1. Two paradigms of data analyses. This figure illustrates the two primary paradigms in data analytics: "prediction" and "causal inference." An example of such "prediction" is the prediction of immunoglobulin A nephropathy progression in pediatric patients based on histological and demographic data, identifying patterns and correlations without addressing the underlying causes. In "causal inference," randomized controlled trials are used to determine whether a specific treatment directly improves patient outcomes, focusing on cause-and-effect relationships.

ference, we can enhance patient outcomes, improve decision-making, and advance healthcare systems. This review explores how AI, particularly predictive modeling and causal inference, can transform nephrology by advancing personalized healthcare. It highlights the strengths and limitations of these approaches, aiming to improve clinical decision-making and outcomes for pediatric kidney patients.

Prediction

Prediction methodology is not designed for inferring causal relationships [15]. Instead, the methodology focuses solely on accurately forecasting outcomes without necessarily understanding the underlying causes that drive these outcomes. In predictive modeling, the primary objective is to develop models that can deliver accurate predictions by identifying patterns and correlations in data, often without regard for whether these correlations represent true causal relationships.

In the context of healthcare, predictive AI algorithms are widely employed to anticipate various outcomes, such as the likelihood of a patient developing a particular disease or the potential for a specific adverse event occurring during treatment. Methods including random forests [16], boosting algorithms [17], support vector machines [18], and deep learning models [19] are commonly employed for these purposes. The algorithms are particularly valuable in scenarios involving large datasets, as they allow for the identification of complex patterns that might not be immediately evident when using traditional statistical methods. For instance, a random forest algorithm can analyze vast amounts of patient data to predict which individuals are at an elevated risk for certain diseases based on patterns identified in their medical histories, demographic data, and other relevant information [20]. Similarly, deep learning models, which are capable of learning intricate representations of data, are increasingly utilized for image recognition tasks, such as identifying abnormalities in medical imaging that might indicate the presence of tumors or other conditions [21].

However, while these predictive models may be highly accurate in forecasting outcomes, they do not inherently provide insights into the reason underlying a particular outcome [22]. These models excel in identifying "what" might happen rather than "why" it happens. For instance, a model might predict that patients with certain characteristics are at high risk for developing diabetes, but it may not clarify which specific factors are causally contributing to the development of the disease. This is because predictive models are fundamentally correlation-based; they identify associations between variables, and the associations do not necessarily imply causation. Furthermore, while predictive models are useful in identifying high-risk groups for certain conditions, they do not facilitate an understanding of the causal pathways that lead to these risks. This limitation is particularly significant in the medical field, where

https://doi.org/10.3339/ckd.24.018

understanding causality is typically crucial for effective intervention. Without knowing the underlying causes, interventions based solely on predictions may not effectively address the root problems [23]. For example, predicting that a patient is at high risk of heart disease based on lifestyle factors is valuable, but without understanding the causal impact of each lifestyle factor, developing targeted strategies for prevention or treatment would be challenging.

Notably, many predictive methods, such as those mentioned above, can be adapted for causal inference purposes under certain conditions. For instance, machine learning models can be employed to estimate causal effects if they are combined with appropriate statistical techniques and study designs, such as instrumental variable analysis or propensity score matching [24,25]. However, such applications require a different set of methodologies and considerations that are beyond the scope of this review. Taken together, while prediction models are indispensable tools for identifying potential risks and outcomes in the medical field, they should not be conflated with causal inference models. The former models focus on correlation and pattern recognition to forecast outcomes, whereas the latter models seek to understand the cause-and-effect relationships that drive those outcomes. Distinguishing between these two approaches is essential to avoid misinterpretation and ensure that the most appropriate methodologies are applied to address specific clinical questions.

Causal inference

Causal inference is a methodology aimed at identifying cause-and-effect relationships, which is crucial for understanding the impact of specific interventions in various domains, including healthcare. In medicine, randomized controlled trials (RCTs) are considered the gold standard for causal inference [26]. The fundamental principle underlying RCTs is the random allocation of participants to different intervention arms, thereby creating comparable groups that can be directly contrasted to estimate the effect of the intervention. This randomization ensures that, on average, the two groups are statistically equivalent concerning both observed and unobserved confounders, thereby allowing for an accurate estimation of causal effects.

Although the RCT design is robust and is considered the most reliable method for causal inference, it is not without limitations. One significant constraint is that RCTs can only estimate the average treatment effect across the study population [27]. This approach does not account for the variability in treatment responses among individuals within the same group, i.e., the methodology cannot estimate the treatment effect for each individual (i.e., the individual treatment effect). In clinical practice, where personalized medicine is becoming increasingly important, understanding how individual patients might respond differently to the same treatment is critical. To address this, interest in estimating heterogeneous treatment effects, which aim to capture the variations in treatment response among subgroups or even at the individual level, is growing [28].

To overcome the limitations of RCTs in estimating heterogeneous treatment effects, various methodologies have been proposed, many of which leverage advanced AI techniques. For example, machine learning models, such as decision trees, random forests, and neural network algorithms, have been adapted to estimate treatment effects across different subgroups defined by covariates.

Local explanation methods

Recent developments in interpretable AI methods have provided tools for understanding and explaining the predictions of these complex models. Several local explanation methods, including Shapley values and Local Interpretable Model-Agnostic Explanations (LIME), have been developed to provide granular insights into how AI models arrive at their predictions [29,30]. Table 1 summarizes the Shapley values and LIME.

Shapley values

Shapley values, which originated from cooperative game theory, provide a solution for fairly distributing the "payout" (in this case, the model's prediction) among different features, based on their contribution to the prediction [29]. When applied to causal inference, Shapley values can help in identifying which variables (or features) are most influential in determining the predicted treatment effect for a particular individual. The strength of this method lies in its ability to offer a theoretically sound approach to feature attribution, ensuring that the contributions of all possible feature combinations are considered. The use of Shapley values is limited by their computational complexity, which can become prohibitively expensive for models with a large number of features or when using extensive datasets. However, the results should be interpreted with caution as they are limited to a specific dataset and they do not imply universal interpretability. Therefore, careful consideration is required when generalizing these findings. As an ex-

Table 1. Summary of the Shapley values and LIME

Aspect	Shapley value	LIME
Origin	Cooperative game theory	Model-agnostic, developed for interpreting black-box models
Approach	Distributes model's prediction based on feature contribution	Approximates complex models with simpler local models
Key strength	Theoretically sound feature attribution considering all combinations	Useful for understanding local decision boundaries
Key limitation	High computational complexity, expensive for large datasets	Relies on linear approximation, which may not hold in non-linear cases
Interpretability	High interpretability	Moderate interpretability, depends on the local model used
Handling non-linear interactions	Can handle non-linear interactions but at high computational cost	Limited handling of non-linear interactions due to linear assumptions

LIME, Local Interpretable Model-Agnostic Explanations.

ample utilizing Shapley values, Oh et al. [20] in their study used this method to identify key demographic factors in predicting coronary calcium scores and selected the most influential factors based on their contribution using Shapley values.

Local Interpretable Model-Agnostic Explanations

Another popular technique to interpret complex models is LIME. The technique involves approximating complex models with simple, interpretable models locally around the prediction of interest [30]. For instance, LIME can fit a linear model around a specific prediction to approximate the behavior of a more complex model in the local vicinity. This approach is particularly useful in understanding the local decision boundaries of black-box models. For example, Li et al. [31] developed and validated a machine learning model in order to predict mortality in critically ill patients with sepsis-associated acute kidney injury; the XGBoost algorithm performed best and they emphasized the use of LIME to interpret individualized predictions and enhance the model's transparency. However, LIME is limited by its reliance on the assumption that a linear model can adequately approximate the complex model locally, which may not always hold true, particularly in cases involving highly non-linear interactions between features.

Local explanation methods have other constraints. They often provide insights that are specific to a particular instance or individual prediction, which may not be generalizable across a broad population [32]. Furthermore, while they can suggest factors that are most influential in a model's prediction, they do not necessarily provide causal explanations or indicate which interventions might lead to desired outcomes [33].

Emulated RCTs in observational studies

Given the limitations and practical constraints of traditional RCTs, particularly in settings where randomization is not feasible or ethical, interest in emulated RCTs, also known as "quasi-experimental designs" or "observational causal inference" has increased [34,35]. These methods aim to replicate the conditions of an RCT using observational data by creating comparable groups that mimic the treatment and control arms in a randomized study.

A key strategy in emulated RCTs is to employ advanced statistical techniques, such as propensity score matching, inverse probability weighting, or regression discontinuity design, to balance the treatment and control groups in terms of observed covariates. This approach attempts to account for confounding variables and simulate the random assignment used in an RCT, thereby enabling a reliable estimation of causal effects from non-randomized data. The potential of using emulated RCTs in the medical field is considerable. They offer a practical solution for evaluating the effectiveness of interventions in real-world settings, where conducting an RCT may be logistically challenging or ethically problematic. For example, using electronic health records, researchers can construct a retrospective cohort study that closely resembles an RCT, making the assessment of treatment effects across different patient populations and clinical settings possible [36,37]. Moreover, with the integration of AI and machine learning, emulated RCTs can be further refined to improve their accuracy and validity. Machine learning algorithms can be employed to identify complex, non-linear relationships between variables and to enhance the precision of propensity score models, thereby reducing residual confounding. Additionally, AI techniques can facilitate

https://doi.org/10.3339/ckd.24.018

the identification of suitable subpopulations for which certain treatments may be particularly effective, thus supporting personalized approaches to healthcare.

Taken together, while traditional RCTs remain the gold standard for causal inference, the emergence of emulated RCTs and advanced AI methodologies provides valuable alternatives for estimating treatment effects in situations where RCTs are not feasible. By leveraging these innovative approaches, deeper insights into causal relationships in healthcare can be achieved, ultimately improving patient outcomes and allowing effective clinical decision-making.

Conclusion

AI has rapidly emerged as a transformative force in healthcare, generating innovative approaches to disease prediction, diagnosis, and patient management. As highlighted in this review, AI applications in the medical field primarily belong to one of two categories: prediction or causal inference. Even though predictive models excel in forecasting outcomes based on historical data, they rely on identifying patterns and correlations without understanding the underlying causative factors, which presents limitations. Causal inference, on the other hand, seeks to address this gap by focusing on the cause-and-effect relationships that drive these outcomes. This inference provides a framework for understanding the mechanisms underlying observed patterns, enabling the development of targeted interventions that can directly influence patient health.

Emulated RCTs leverage observational data to mimic the conditions of randomized trials, providing a promising avenue for causal inference in scenarios where traditional RCTs are impractical. Combined with sophisticated AI and machine learning algorithms, these approaches can identify complex, non-linear relationships between variables, enhance the precision of causal estimates, and reduce residual confounding.

The convergence of prediction and causal inference in AI holds great promise for the future of healthcare. By integrating these approaches, we can better understand why outcomes happen and how to intervene effectively, leading to more personalized patient care. However, if prediction and causal inference are not properly distinguished, significant issues can arise. The future of medicine lies in harnessing AI's potential not only for prediction and diagnosis but also for understanding and transforming patient care for the better.

Conflicts of interest

No potential conflict of interest relevant to this article was reported.

Funding

None.

Author contributions

All the work was done by TRO.

References

1. Alanazi R. Identification and prediction of chronic diseases using machine learning approach. J Healthc Eng 2022;2022:2826127.
2. Kumar A, Satyanarayana Reddy SS, Mahommad GB, Khan B, Sharma R. Smart healthcare: disease prediction using the cuckoo-enabled deep classifier in IoT framework. Sci Program 2022;2022:2090681.
3. Talukdar J, Singh TP. Early prediction of cardiovascular disease using artificial neural network. Paladyn J Behav Robot 2023;14:20220107.
4. Tomasev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature 2019;572:116-9.
5. Kavitha C, Mani V, Srividhya SR, Khalaf OI, Tavera Romero CA. Early-stage Alzheimer's disease prediction using machine learning models. Front Public Health 2022;10:853294.
6. Basu S, Sussman JB, Hayward RA. Detecting heterogeneous treatment effects to guide personalized blood pressure treatment: a modeling study of randomized clinical trials. Ann Intern Med 2017;166:354-60.
7. Govindaraj M, Asha V, Saju B, Sagar M, Rahul. Machine learning algorithms for disease prediction analysis. In: 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). Tirunelveli, India; 2023. p. 879-88.
8. Verma VK, Lin WY. A machine learning-based predictive model for 30-day hospital readmission prediction for COPD patients. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Toronto, Canada; 2020. p. 994-9.
9. Anurag, Vyas N, Sharma V, Balla D. Chronic kidney disease prediction using robust approach in machine learning. In: 2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT). Dehradun, India; 2023. p. 1-5.
10. Pearl J. Causal inference in statistics: an overview. Stat Surv 2009;

3:96-146.

11. Prosperi M, Guo Y, Sperrin M, Koopman JS, Min JS, He X, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nat Mach Intell 2020;2:369-75.
12. Shen X, Ma S, Vemuri P, Castro MR, Caraballo PJ, Simon GJ. A novel method for causal structure discovery from EHR data and its application to type-2 diabetes mellitus. Sci Rep 2021;11:21025.
13. Sanchez P, Voisey JP, Xia T, Watson HI, O'Neil AQ, Tsaftaris SA. Causal machine learning for healthcare and precision medicine. R Soc Open Sci 2022;9:220638.
14. Phillips DP, Liu GC, Kwok K, Jarvinen JR, Zhang W, Abramson IS. The Hound of the Baskervilles effect: natural experiment on the influence of psychological stress on timing of death. BMJ 2001;323:1443-6.
15. Arif S, MacNeil MA. Predictive models aren't for causal inference. Ecol Lett 2022;25:1741-5.
16. Breiman L. Random forests. Mach Learn 2001;45:5-32.
17. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 1997;55:119-39.
18. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273-97.
19. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86:2278-324.
20. Oh TR, Song SH, Choi HS, Suh SH, Kim CS, Jung JY, et al. Predictive model for high coronary artery calcium score in young patients with non-dialysis chronic kidney disease. J Pers Med 2021;11:1372.
21. Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. Clin Cancer Res 2019;25:3266-75.
22. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. Entropy (Basel) 2020;23:18.
23. Carloni G, Berti A, Colantonio S. The role of causality in explainable artificial intelligence. arXiv [Preprint] 2023 Sep 18. https:// doi.org/10.48550/arXiv.2309.09901
24. Ichimura H, Taber C. Propensity-score matching with instrumental variables. Am Econ Rev 2001;91:119-24.
25. Heckman J, Navarro-Lozano S. Using matching, instrumental variables, and control functions to estimate economic choice models. Rev Econ Stat 2004;86:30-57.
26. Hariton E, Locascio JJ. Randomised controlled trials: the gold standard for effectiveness research: Study design: randomised controlled trials. BJOG 2018;125:1716.
27. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Soc Sci Med 2018;210:2-21.
28. Rekkas A, Paulus JK, Raman G, Wong JB, Steyerberg EW, Rijnbeek PR, et al. Predictive approaches to heterogeneous treatment effects: a scoping review. BMC Med Res Methodol 2020;20:264.
29. Messalas A, Kanellopoulos Y, Makris C. Model-agnostic interpretability with Shapley values. In: 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA). Patras, Greece; 2019. p. 1-7.
30. Zafar MR, Khan NM. DLIME: a deterministic Local Interpretable Model-Agnostic Explanations approach for computer-aided diagnosis systems. arXiv [Preprint] 2019 Jun 24. https://doi.org/10.48550/arXiv.1906.10263
31. Li X, Wu R, Zhao W, Shi R, Zhu Y, Wang Z, et al. Machine learning algorithm to predict mortality in critically ill patients with sepsis-associated acute kidney injury. Sci Rep 2023;13:5223.
32. Raghavan S, Josey K, Bahn G, Reda D, Basu S, Berkowitz SA, et al. Generalizability of heterogeneous treatment effects based on causal forests applied to two randomized clinical trials of intensive glycemic control. Ann Epidemiol 2022;65:101-8.
33. Pichler M, Hartig F. Can predictive models be used for causal inference? arXiv [Preprint] 2023 Jun 18. https://doi.org/10.48550/arXiv.2306.10551
34. Kutcher SA, Brophy JM, Banack HR, Kaufman JS, Samuel M. Emulating a randomised controlled trial with observational data: an introduction to the target trial framework. Can J Cardiol 2021;37:1365-77.
35. Gianicolo EA, Eichler M, Muensterer O, Strauch K, Blettner M. Methods for evaluating causality in observational studies. Dtsch Arztebl Int 2020;116:101-7.
36. Rasouli B, Chubak J, Floyd JS, Psaty BM, Nguyen M, Walker RL, et al. Combining high quality data with rigorous methods: emulation of a target trial using electronic health records and a nested case-control design. BMJ 2023;383:e072346.
37. Sengupta S, Ntambwe I, Tan K, Liang Q, Paulucci D, Castellanos E, et al. Emulating randomized controlled trials with hybrid control arms in oncology: a case study. Clin Pharmacol Ther 2023;113:867-77.

https://doi.org/10.3339/ckd.24.018