


LETTER TO THE EDITOR

Open Access



The application of time-to-event analysis in machine learning prognostic models

Zi-He Peng^{1,2}, Zhi-Xin Huang^{1,2}, Juan-Hua Tian¹, Tie Chong^{1*} and Zhao-Lun Li^{1*} 

To the editor,

Artificial intelligence (AI), with its ambition to emulate human intelligence within machines, has emerged as a transformative force across various domains. Within the realm of modern medicine, the digitization of practices such as electronic medical records has ushered in new opportunities for the integration of machine learning (ML) methodologies. These innovations find application in diverse contexts, from AI-assisted pathology assessments to the ML-driven analysis of qualitative interviews and medical records, unearthing intricate themes and underlying patterns. In the clinical sphere, ML frequently directs its focus toward enhancing predictive capabilities, harnessing the potential of commonplace and readily accessible variables to refine prognostic accuracy. It's worth noting that while many ML analyses focus on classification issues and the creation of diagnostic models, in the medical field, the prevalent approach involves using survival analysis to develop prognostic models.

Survival analysis, an intricate statistical method, is designed to unravel the intricate correlations between covariates and the temporal occurrences of events. Unlike conventional classification paradigms, survival analysis confronts the intricacies engendered by partially observed data, often stemming from censoring.

Within the realm of clinical inquiries, patient records manifest in distinct categories: those that remain uncensored, thereby divulging precise event timings, and those that exist as right-censored, withholding event timings beyond the study's temporal scope. This distinctive attribute mandates the utilization of specialized models adept at accommodating the complexities inherent to such data structures, thus emerging as a pivotal facet within the realm of survival analysis methodology.

Regrettably, it is observed that many recently published articles have erred by simplistically transforming outcomes into categorical variables and utilizing ML classification techniques to formulate prognostic models [1, 2]. These endeavors have been undertaken without due consideration for the impact of censored data on the model's fidelity. A systematic review uncovered that among 11 studies crafting 24 models for survival outcomes, merely ten models explicitly took into account censored observations, of which seven were built upon the framework of Cox regression [3]. This implies that only three ML models are considered censored data. These studies usually exclude patients who survive but for shorter than a specific date, after which several date-specific models are built (e.g., 3-year, 5-year). To employ a straightforward analogy, where a traditional statistical model, specifically Cox proportional hazard regression, should have been adopted to construct prognostic models predicting survival at 3-year and 5-year intervals, logistic regression was employed to create two categorical models. As underscored by PROBAST (Prediction model risk of Bias ASsessment Tool), the exclusion of censored participants through simplistic logistic regression models is deemed unsuitable [4]. The utilization of an erroneous logistic regression methodology results in a selected dataset

*Correspondence:

Tie Chong

chongtie@126.com

Zhao-Lun Li

oliverlee0615@163.com

¹ Department of Urology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710004, China

² Health Science Center, Xi'an Jiaotong University, Xi'an 710061, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

containing fewer individuals lacking the outcome, thus introducing bias into predicted risks due to the overrepresentation of those with the outcome [4]. The time-to-event analysis provides an effective means of addressing these censored observations. Contrary to the notion that there are no ML algorithm packages for conducting survival analysis, there is indeed a Python module named “scikit-survival” designed for this purpose. It is developed on top of scikit-learn and can be found at <https://scikit-survival.readthedocs.io/en/latest/index.html> (accessed on October 30, 2023) [5]. This module enables the incorporation of survival analysis within the capabilities of scikit-learn. We highly advise employing “scikit-survival” for the development of ML prognostic models.

When developing ML prognostic models, it is strongly advised to employ survival analysis techniques such as “scikit-survival” to appropriately handle censored observations. Simply excluding or categorizing censored cases using logistic regression is inappropriate and introduces bias. Overall, ignoring censoring and using inaccurate evaluation metrics can severely compromise the validity of machine learning-based prognostic models. Careful consideration of censoring and time-to-event analysis principles is warranted.

Abbreviations

AI	Artificial intelligence
ML	Machine learning
PROBAST	Prediction model risk of Bias ASsessment Tool

Acknowledgements

None.

Author contributions

Conceptualization: ZHP, TC and ZLL; writing—original draft: ZHP; writing—review and editing: ZXH and JHT; project administration: TC and ZLL. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

None.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Karabacak M, Jagtiani P, Carrasquilla A, Germano IM, Margetis K. Prognosis individualized: survival predictions for WHO grade II and III gliomas with a machine learning-based web application. *NPJ Digit Med.* 2023;6(1):200.
2. Li C, Liu M, Zhang Y, Wang Y, Li J, Sun S, Liu X, Wu H, Feng C, Yao P, et al. Novel models by machine learning to predict prognosis of breast cancer brain metastases. *J Transl Med.* 2023;21(1):404.
3. Dhiman P, Ma J, Andaur Navarro CL, Speich B, Bullock G, Damen JAA, Hooft L, Kirtley S, Riley RD, Van Calster B, et al. Methodological conduct of prognostic prediction models developed using machine learning in oncology: a systematic review. *BMC Med Res Methodol.* 2022;22(1):101.
4. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med.* 2019;170(1):W1–33.
5. Polsterl S. scikit-survival: a library for Time-to-event analysis built on top of scikit-learn. *J Mach Learn Res.* 2020;21:8747.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.