



## Machine learning in liver surgery: Benefits and pitfalls

Rafael Calleja, Manuel Durán, María Dolores Ayllón, Ruben Ciria, Javier Briceño

**Specialty type:** Medicine, research and experimental

**Provenance and peer review:**

Unsolicited article; Externally peer reviewed.

**Peer-review model:** Single blind

**Peer-review report's scientific quality classification**

Grade A (Excellent): A  
Grade B (Very good): 0  
Grade C (Good): 0  
Grade D (Fair): 0  
Grade E (Poor): 0

**P-Reviewer:** Liu S, China

**Received:** December 20, 2023

**Revised:** February 8, 2024

**Accepted:** March 29, 2024

**Published online:** April 26, 2024



**Rafael Calleja, Manuel Durán, María Dolores Ayllón, Ruben Ciria, Javier Briceño**, Hepatobiliary Surgery and Liver Transplantation Unit, Hospital Universitario Reina Sofía, Maimonides Biomedical Research Institute of Cordoba, Córdoba 14004, Spain

**Corresponding author:** Rafael Calleja, MD, Research Associate, Surgeon, Hepatobiliary Surgery and Liver Transplantation Unit, Hospital Universitario Reina Sofía, Maimonides Biomedical Research Institute of Cordoba, Avenida Menéndez Pidal s/n, Córdoba 14004, Spain. [h12calor@gmail.com](mailto:h12calor@gmail.com)

### Abstract

The application of machine learning (ML) algorithms in various fields of hepatology is an issue of interest. However, we must be cautious with the results. In this letter, based on a published ML prediction model for acute kidney injury after liver surgery, we discuss some limitations of ML models and how they may be addressed in the future. Although the future faces significant challenges, it also holds a great potential.

**Key Words:** Machine learning; Liver surgery; Artificial intelligence; Random forest; Prediction model

©The Author(s) 2024. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core Tip:** Artificial intelligence is trending topic in healthcare research. Machine learning classifiers have been explored in the field of liver surgery and liver transplantation. However, despite of promising results, a real applicability is limited by several factors.

**Citation:** Calleja R, Durán M, Ayllón MD, Ciria R, Briceño J. Machine learning in liver surgery: Benefits and pitfalls. *World J Clin Cases* 2024; 12(12): 2134-2137

**URL:** <https://www.wjgnet.com/2307-8960/full/v12/i12/2134.htm>

**DOI:** <https://dx.doi.org/10.12998/wjcc.v12.i12.2134>

### TO THE EDITOR

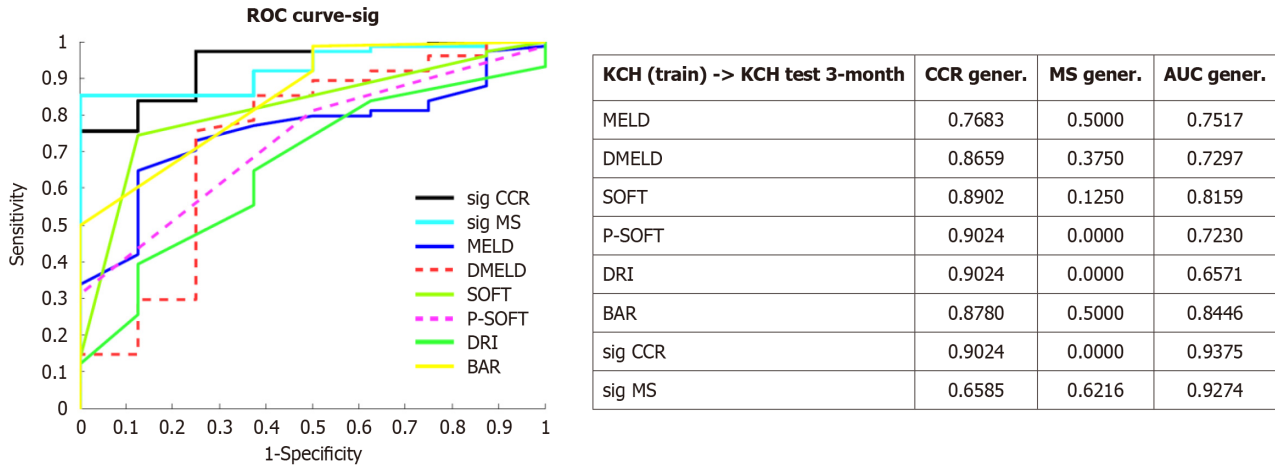
We read with interest the retrospective study by Dong *et al*[1] that developed a machine learning (ML) prediction model for acute kidney injury (AKI) following liver

resection (LR). We thank the authors for their work and contribution in this field. LR is the first-line treatment of various liver lesions. However, the reported incidence of AKI after LR ranges from 10% to 15% [2], significantly impacting patient morbidity and mortality. Hence, identifying factors that may lead to the development of AKI is relevant. Dong *et al* [1] explored the potential contribution of ML classifiers to this issue.

The authors analyzed a retrospective cohort of 2450 patients and trained and validated four ML classifiers (logistic regression, random forest, support vector machine, extreme gradient boosting, and decision tree). The training methodology (10-fold cross-validation) and validation (a holdout technique with 30% patterns) were adequate. Random forest exhibited the highest performance [area under the curve (AUC) = 0.92] among the classifiers. Although the results were satisfactory, certain considerations must be addressed.

First, the rate of missing values should be reported because it can affect model training, subsequently affecting model performance and generalizability. Hence, random forest classifiers are the best algorithms for a significant rate of missing values [3]. Conversely, if this rate is low, artificial neural networks (ANNs) could offer promising dataset results. Second, several factors reported in the literature are associated with AKI after LR, such as major hepatectomy, surgery duration, hepatojejunostomy, increased Model for End-Stage Liver Disease score (MELD), and blood transfusion [2,4-8]. Among these factors, only surgery duration was included in the baseline characteristics. The inclusion of these variables may have increased the robustness of the model. Finally, performing external validation is challenging. Differences between the training and external validation cohorts may impact model accuracy. Therefore, a prospective validation may be an alternative.

Some recent studies in ML applications ranges from protein structure prediction or COVID-19 diagnosis from X-ray images to optimizing donor-recipient matching to reduce waitlist mortality or improve post-transplant outcomes [9-11]. Our experience in the field of ML in liver surgery started from liver transplantation and efforts primarily focused on improving donor-recipient matching. Using graft survival as the endpoint, we developed an ANN model that achieved an AUC of approximately 0.8212 [12]. This method was validated in an external cohort and improved AUC by 15% [13]. This ANN was integrated into a rule system with the MELD score to prioritize graft allocation. Although this method was explored in the United Network of Organ Sharing database, limited results were obtained because of a significant proportion of missing values were found [14]. Dong *et al* [1] found that the model performance was better than the current scores for AKI prediction. Similarly, we reported the difference of ML models that outperformed traditional scores, such as MELD, Survival Outcomes Following Liver Transplantation score, Donor Risk Index score, and Balance of Risk score (Figure 1). In medicine, certain variables do not necessarily have to assume a linear relationship. Hence, ML models are superior to statistical methods (linear regression), from which most of these scores are derived [15]. However, these findings may be attributed to model overtraining; therefore, validation is required.



**Figure 1 External validation of artificial neural network models [13].** The performance obtained by these models is compared to other published score in terms of area under curve. A receiver operating characteristic curve depicts these metrics. Artificial neural network models based on the concept of minimum sensitivity and correct classification rate are represented such as sig minimum sensitivity and sig correct classification rate respectively. These models outperformed other traditional scores such as Model for End-Stage Liver Disease, Model for End-Stage Liver Disease score excluding exception points and donor age, Survival Outcomes Following Liver Transplantation, Preallocation Survival Outcomes Following Liver Transplantation, Donor Risk Index or Balance of Risk. CCR: Correct classification rate; MS: Minimum sensitivity; MELD: Model for End-Stage Liver Disease score; DMELD: Model for End-Stage Liver Disease score excluding exception points and donor age; SOFT: Survival Outcomes Following Liver Transplantation score; P-SOFT: Preallocation Survival Outcomes Following Liver Transplantation score; DRI: Donor Risk Index score; BAR: Balance of Risk score; AUC: Area under curve; KCH: Kings College Hospital; ROC: Receiver operating characteristic. Citation: Ayllón MD, Ciria R, Cruz-Ramírez M, Pérez-Ortiz M, Gómez I, Valente R, O’Grady J, de la Mata M, Hervás-Martínez C, Heaton ND, Briceño J. Validation of artificial neural networks as a methodology for donor-recipient matching for liver transplantation. *Liver Transpl* 2018; **24**: 192-203. Copyright© The Authors 2018. Published by Wolters Kluwer Health, Inc.

The most significant lesson learned from using these models is their high dependency on the datasets on which they were trained. This issue affects the practical applicability. Retrospective data, external validation, the “black box issues” in ANN, and data-protection policies are considered significant contributing factors. To overcome these barriers, better

data-handling policies are needed. Applicability relies on the clinicians' confidence in using these models. Therefore, if external validation is impossible (region-specific rather than universal models), prospective validation should be considered. Moreover, the databases must be updated regularly to reinforce the learning of these models. Clinical scenarios are dynamic, and models must change accordingly.

Recently, interest in artificial intelligence and ML has increased. They can handle large amounts of data quickly and yield accurate results. However, we must note the limitations of these models and address them to achieve a real integration.

## FOOTNOTES

**Author contributions:** Calleja R and Durán M designed and wrote this letter; Ayllón MD, Ciria R and Briceño J performed the group research mentioned in this letter; All authors have read and approve the final manuscript.

**Conflict-of-interest statement:** The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this manuscript. This research has not received any financial support.

**Open-Access:** This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

**Country/Territory of origin:** Spain

**ORCID number:** Rafael Calleja 0000-0003-0048-0756; Manuel Durán 0000-0003-1161-2195; María Dolores Ayllón 0000-0003-2493-3756; Ruben Ciria 0000-0002-7839-2329; Javier Briceño 0000-0001-7027-7898.

**S-Editor:** Che XX

**L-Editor:** A

**P-Editor:** Zheng XM

## REFERENCES

- Dong JF, Xue Q, Chen T, Zhao YY, Fu H, Guo WY, Ji JS. Machine learning approach to predict acute kidney injury after liver surgery. *World J Clin Cases* 2021; **9**: 11255-11264 [PMID: 35071556 DOI: 10.12998/wjcc.v9.i36.11255]
- Lim C, Audureau E, Salloum C, Levesque E, Lahat E, Merle JC, Compagnon P, Dhonneur G, Feray C, Azoulay D. Acute kidney injury following hepatectomy for hepatocellular carcinoma: incidence, risk factors and prognostic value. *HPB (Oxford)* 2016; **18**: 540-548 [PMID: 27317959 DOI: 10.1016/j.hpb.2016.04.004]
- Calleja Lozano R, Hervás Martínez C, Briceño Delgado FJ. Crossroads in Liver Transplantation: Is Artificial Intelligence the Key to Donor-Recipient Matching? *Medicina (Kaunas)* 2022; **58** [PMID: 36556945 DOI: 10.3390/medicina58121743]
- Kambakamba P, Slankamenac K, Tschuor C, Kron P, Wirsching A, Maurer K, Petrowsky H, Clavien PA, Lesurtel M. Epidural analgesia and perioperative kidney function after major liver resection. *Br J Surg* 2015; **102**: 805-812 [PMID: 25877255 DOI: 10.1002/bjs.9810]
- Peres LA, Bredt LC, Cipriani RF. Acute renal injury after partial hepatectomy. *World J Hepatol* 2016; **8**: 891-901 [PMID: 27478539 DOI: 10.4254/wjh.v8.i21.891]
- Reese T, Kröger F, Makridis G, Drexler R, Jusufi M, Schneider M, Brüning R, von Rittberg Y, Wagner KC, Oldhafer KJ. Impact of acute kidney injury after extended liver resections. *HPB (Oxford)* 2021; **23**: 1000-1007 [PMID: 33191106 DOI: 10.1016/j.hpb.2020.10.015]
- Slankamenac K, Beck-Schimmer B, Breitenstein S, Puhan MA, Clavien PA. Novel prediction score including pre- and intraoperative parameters best predicts acute kidney injury after liver surgery. *World J Surg* 2013; **37**: 2618-2628 [PMID: 23959337 DOI: 10.1007/s00268-013-2159-6]
- Tomozawa A, Ishikawa S, Shiota N, Cholvisudhi P, Makita K. Perioperative risk factors for acute kidney injury after liver resection surgery: an historical cohort study. *Can J Anaesth* 2015; **62**: 753-761 [PMID: 25925634 DOI: 10.1007/s12630-015-0397-9]
- Bhat M, Rabindranath M, Chara BS, Simonetto DA. Artificial intelligence, machine learning, and deep learning in liver transplantation. *J Hepatol* 2023; **78**: 1216-1233 [PMID: 37208107 DOI: 10.1016/j.jhep.2023.01.006]
- Cai C, Gou B, Khishe M, Mohammadi M, Rashidi S, Moradpour R, Mirjalili S. Improved deep convolutional neural networks using chimp optimization algorithm for Covid19 diagnosis from the X-ray images. *Expert Syst Appl* 2023; **213**: 119206 [PMID: 36348736 DOI: 10.1016/j.eswa.2022.119206]
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021; **596**: 583-589 [PMID: 34265844 DOI: 10.1038/s41586-021-03819-2]
- Briceño J, Cruz-Ramírez M, Prieto M, Navasa M, Ortiz de Urbina J, Orti R, Gómez-Bravo MÁ, Otero A, Varo E, Tomás S, Clemente G, Bañares R, Bárcena R, Cuervas-Mons V, Solórzano G, Vinaixa C, Rubín A, Colmenero J, Valdivieso A, Ciria R, Hervás-Martínez C, de la Mata M. Use of artificial intelligence as an innovative donor-recipient matching model for liver transplantation: results from a multicenter Spanish study. *J Hepatol* 2014; **61**: 1020-1028 [PMID: 24905493 DOI: 10.1016/j.jhep.2014.05.039]

- 13 **Ayllón MD**, Ciria R, Cruz-Ramírez M, Pérez-Ortiz M, Gómez I, Valente R, O'Grady J, de la Mata M, Hervás-Martínez C, Heaton ND, Briceño J. Validation of artificial neural networks as a methodology for donor-recipient matching for liver transplantation. *Liver Transpl* 2018; **24**: 192-203 [PMID: [28921876](#) DOI: [10.1002/lt.24870](#)]
- 14 **Guijo-Rubio D**, Briceño J, Gutiérrez PA, Ayllón MD, Ciria R, Hervás-Martínez C. Statistical methods versus machine learning techniques for donor-recipient matching in liver transplantation. *PLoS One* 2021; **16**: e0252068 [PMID: [34019601](#) DOI: [10.1371/journal.pone.0252068](#)]
- 15 **Briceño J**, Calleja R, Hervás C. Artificial intelligence and liver transplantation: Looking for the best donor-recipient pairing. *Hepatobiliary Pancreat Dis Int* 2022; **21**: 347-353 [PMID: [35321836](#) DOI: [10.1016/j.hbpd.2022.03.001](#)]



Published by **Baishideng Publishing Group Inc**  
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

**Telephone:** +1-925-3991568

**E-mail:** [office@baishideng.com](mailto:office@baishideng.com)

**Help Desk:** <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

