Letter to the Editor

JOURNAL OF HEPATOLOGY

Revealing bias in feature importance through PLS-DA: A critical examination of machine learning applications in chronic liver disease

To the Editor:

To assess clinical factors, perform genotyping of known variants, and conduct comprehensive metabolic phenotyping aimed at characterizing the regression of fibrosis in patients with compensated advanced chronic liver disease, Mendoza *et al.* employed a partial least squares discriminant analysis (PLS-DA) model. This methodology sought to evaluate the importance of individual variables in distinguishing between patients who experienced regression of fibrosis and those who did not.[1]

While Mendoza *et al.* emphasize the accuracy of their machine learning predictions, this paper raises critical concerns about their application of the PLS-DA model in assessing feature importance. Our analysis suggests a fundamental misunderstanding of key principles underlying machine learning methodology. While the primary objective of machine learning is to achieve precise predictions of the target variable, the computation of feature importances is intended to clarify the genuine relationships between the target and the features. However, the model-specific nature of these importance measures can introduce biases, potentially leading to misleading interpretations. More than 100 peer-reviewed publications have thoroughly examined the issue of bias in feature importance derived from machine learning models, including PLS-DA.[2–5] Furthermore, achieving high cross-validation accuracy for target prediction does not necessarily validate the accuracy of the associated feature importance.

The implications of biased feature importances are far from trivial; they can significantly influence clinical decision-making and the interpretation of research findings. Misleading feature importances may distort our understanding of which factors truly influence outcomes, potentially derailing further research initiatives and skewing treatment strategies. Given the intricate and multifaceted nature of chronic liver disease, reliance on biased results could lead to ineffective or even harmful clinical practices.

Thus, it is critical to approach the analysis of feature importances with caution, acknowledging the inherent potential for bias in many machine learning models. We strongly advocate for the use of robust, bias-free statistical techniques, such as Spearman's correlation combined with *p* values, to ascertain genuine relationships between features and clinical outcomes.[6–9] By adhering to these rigorous analytical standards, researchers can ensure their findings contribute meaningful insights to the understanding and treatment of complex diseases like chronic liver disease.

Different machine learning models employ varying methodologies for generating feature importance, resulting in distinct inherent biases associated with each model. Consequently, reliance on feature importances as definitive indicators of true associations between variables can lead to erroneous conclusions. It is crucial to recognize that feature importances derived from models like PLS-DA should not be treated as conclusive relationships.

To establish authentic relationships between the target and features, three key elements must be considered to ensure the reliability of the relationship values: thorough assessment of data distribution, examination of statistical relationships, and determination of statistical significance through *p* values. Based on these elements, careful consideration must also be given to the choice of model – whether linear or non-linear – and the selection of parametric or non-parametric approaches. Since the PLS-DA model is inherently a linear and parametric method, researchers must diligently evaluate the data for linear relationships and parametric assumptions prior to its application. However, Mendoza *et al.* appear to have overlooked these critical checks, which ultimately compromises the integrity of their findings.

In conclusion, we underscore the importance of employing sound statistical rigor when interpreting relationships between variables in clinical research. A more nuanced approach that prioritizes verification of assumptions and consideration of alternative methods will yield more reliable and meaningful insights into the associations that underpin complex diseases such as chronic liver disease, ultimately improving patient management.

Yoshiyasu Takefuji*
Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan
*Address: Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan; Tel.: +81-3-6865-7681.
E-mail address: takefuji@keio.jp

Received 2 December 2024; Accepted 8 December 2024; Available online xxx
https://doi.org/10.1016/j.jhep.2024.12.021

Financial support

The author did not receive any financial support to produce this manuscript.

Conflict of interest

The author declares no conflict of interest.

Please refer to the accompanying ICMJE disclosure form for further details.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2024.12.021.

References

[1] Mendoza YP, Tsouka S, Semmler G, et al. Metabolic phenotyping of patients with advanced chronic liver disease for better characterization of cirrhosis regression. J Hepatol 2024;81(6):983–994. https://doi.org/10.1016/j.jhep.2024.06.028.

[2] Mendez KM, Broadhurst DI, Reinke SN. Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks. Metabolomics 2020;16(2):17. https://doi.org/10.1007/s11306-020-1640-0. Published 2020 Jan 21.

[3] Bifarin OO. Interpretable machine learning with tree-based shapley additive explanations: application to metabolomics datasets for binary classification. PLoS One 2023;18(5):e0284315. https://doi.org/10.1371/journal.pone.0284315. Published 2023 May 4.

[4] Shi L, Westerhuis JA, Rosén J, et al. Variable selection and validation in multivariate modelling. Bioinformatics 2019;35(6):972–980. https://doi.org/10.1093/bioinformatics/bty710.

[5] van Reenen M, Reinecke CJ, Westerhuis JA, et al. Variable selection for binary classification using error rate p-values applied to metabolomics data. BMC Bioinformatics 2016;17:33. https://doi.org/10.1186/s12859-015-0867-7. Published 2016 Jan 14.

[6] Yu H, Hutson AD. A robust Spearman correlation coefficient permutation test. Commun Stat Theor Methods 2024;53(6):2141–2153. https://doi.org/10.1080/03610926.2022.2121144.

[7] Rovetta A. Raiders of the lost correlation: a guide on using pearson and spearman coefficients to detect hidden correlations in medical sciences. Cureus 2020;12(11):e11794. https://doi.org/10.7759/cureus.11794. Published 2020 Nov 30.

[8] Liu Q, Li C, Wanga V, et al. Covariate-adjusted Spearman's rank correlation with probability-scale residuals. Biometrics 2018;74(2):595–605. https://doi.org/10.1111/biom.12812.

[9] Eden SK, Li C, Shepherd BE. Nonparametric estimation of Spearman's rank correlation with bivariate survival data. Biometrics 2022;78(2):421–434. https://doi.org/10.1111/biom.13453.