

# Addressing Bias in Feature Importance: A Hybrid Approach for Risk Prediction in Prognostic Survival Models

## TO THE EDITOR:

To deliver reliable prognostic survival guidance for new patients, Zhang et al implemented a hybrid feature selection technique that integrated both filter and wrapper methods. Their framework used 12 machine learning models, including random forest, gradient boosting, support vector machines, and other models, identifying 10 critical features for risk prediction through SHAP.<sup>1</sup> The validity of their results was assessed using SHAP, which, despite relying on machine learning models, underscores that the accuracy of feature selection is essential for establishing their validity.

However, the reliance on machine learning models for feature selection introduces inherent biases that may lead to incorrect conclusions. Instead of depending on potentially biased feature importances, it is essential to use true associations between the target variable and features. This paper discusses the biases induced by random survival forests and advocates for the use of robust statistical methods, such as Chi-squared tests with *P* values or Spearman correlation with *P* values, to determine genuine associations.

The random forest algorithm has been adapted for various clinical applications, including random survival forests.<sup>2</sup> Consequently, both random survival forests and standard random forests exhibit similar bias behaviors because of their reliance on ensemble decision trees. This paper investigates the mechanisms by which these ensemble decision trees introduce biases in feature importance assessments. Consequently, feature selection on the basis of random survival forests can lead to misleading conclusions.

Researchers must recognize that the primary goal of machine learning is to accurately predict the target variable, while feature importances derived from machine learning models are ultimately by-products of this predictive process. Given the model-specific nature of machine learning, these feature importances are inherently susceptible to biases. As such, a nuanced understanding of these biases is essential for drawing valid conclusions from model outputs and ensuring the robustness of findings in clinical applications. This paper recommends the use of true associations using statistical methods such as Chi-squared tests with *P* values<sup>3,4</sup> or Spearman correlation with *P* values.<sup>5,6</sup>

This paper not only summarizes the factors contributing to feature importance biases induced by decision trees<sup>7-10</sup> but also advocates for the use of true associations between the target variable and features. By using more accurate

statistical methods, researchers can avoid these pitfalls and enhance the robustness of their findings.

Ensemble decision trees, such as random forests including random survival forests and gradient boosting machines, are widely used in machine learning because of their high predictive performance. However, one notable issue with these methods is their propensity to generate biased feature importance scores. This bias primarily arises from the way categorical variables are handled in decision trees.

When using ensemble methods, features are typically assessed on the basis of their ability to improve model performance. Each tree in the ensemble may split on a feature multiple times, and if that feature is categorical with many unique values, it can lead to inflated importance scores. In scenarios where a feature offers numerous specific categories, the tree can easily gain multiple boosts by leveraging these splits. As a result, this feature may seem disproportionately influential compared with others that do not benefit from as many splits.

Another source of bias is multicollinearity among features. When multiple correlated features exist, decision trees often favor one over the other during splitting. This favoritism can lead to misleading conclusions about which features are truly important for predicting the target variable. In essence, one feature may absorb the importance that would otherwise be attributed to its correlated counterparts, creating an illusion of singular dominance.

Additionally, overfitting is a significant concern with ensemble decision trees. As models become more complex and learn intricate patterns in the training data, they may fit noise rather than the true underlying relationships. This overfitting can result in overly optimistic feature importance scores, as the model may attribute significant importance to features that do not generalize well to new data. Furthermore, the inherent randomness in ensemble methods can exacerbate the problem. As trees are built using random subsets of data and features, certain features might gain prominence simply because they happen to be selected in numerous ensembles. This stochastic nature can skew the importance rankings, leading to a preference for features that may not genuinely contribute to the overall model performance.

Ensemble decision trees, such as random forests and gradient boosting machines, offer high predictive performance but often produce biased feature importance scores. This

Corresponding author: Yoshiyasu Takefuji, PhD; e-mail: takefuji@keio.jp.

bias arises primarily from how categorical variables and multicollinearity are handled within decision trees. When split across multiple unique categories, certain features can seem artificially more influential than they are, while others may absorb importance because of correlation. Additionally, overfitting and the inherent randomness in ensemble methods can further distort feature significance. Our analysis highlights the necessity of using robust statistical methods, notably Chi-squared tests and Spearman correlation, to accurately capture genuine associations between features and the target variable. Adopting these practices will improve the validity of findings and enhance predictive accuracy in clinical settings.

**Yoshiyasu Takefuji, PhD** 

*Faculty of Data Science, Musashino University, Tokyo, Japan*

## AUTHOR'S DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by the author of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/po/author-center](http://ascopubs.org/po/author-center).

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://Open Payments)).

No potential conflicts of interest were reported.

## REFERENCES

- Zhang S, Zhang G, Wang M, et al: Artificial intelligence hybrid survival assessment system for robot-assisted proctectomy: A retrospective cohort study. *JCO Precis Oncol* **10.1200/PO.24.00089**
- Bohannon ZS, Coffman F, Mitrofanova A: Random survival forest model identifies novel biomarkers of event-free survival in high-risk pediatric acute lymphoblastic leukemia. *Comput Struct Biotechnol J* **20:583-597**, 2022
- Tadesse AW, Ayano G, Dachew BA, et al: Exposure to maternal cannabis use disorder and risk of autism spectrum disorder in offspring: A data linkage cohort study. *Psychiatry Res* **337:115971**, 2024
- Swami V, Maijano C, Morin AJS: A guide to exploratory structural equation modeling (ESEM) and bifactor-ESEM in body image research. *Body Image* **47:101641**, 2023
- Feiler MO, Kulick ER, Sinclair K, et al: Toxic metals and pediatric clinical immune dysfunction: A systematic review of the epidemiological evidence. *Sci Total Environ* **927:172303**, 2024
- Lu X, Vano Y-A, Su X, et al: Silencing of genes by promoter hypermethylation shapes tumor microenvironment and resistance to immunotherapy in clear-cell renal cell carcinomas. *Cell Rep Med* **4:101287**, 2023
- Chanderraj R, Brown CA, Hinkle K, et al: Gut microbiota predict enterococcus expansion but not vancomycin-resistant enterococcus acquisition. *mSphere* **5:e00537-20**, 2020
- Qian H, Wang B, Yuan M, et al: Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. *Expert Syst Appl* **190:116202**, 2022
- Chen V, Yang M, Cui W, et al: Applying interpretable machine learning in computational biology—pitfalls, recommendations and opportunities for new developments. *Nat Methods* **21:1454-1461**, 2024
- Aniceto N, Bonifácio VDB, Guedes RC, et al: Exploring the chemical space of urease inhibitors to extract meaningful trends and drivers of activity. *J Chem Inf Model* **62:3535-3550**, 2022

DOI: <https://doi.org/10.1200/PO-24-00785>; Published at [ascopubs.org/journal/po](http://ascopubs.org/journal/po) on XXXX, 2024.