Dear Editor,

The Current Opinion article by Bullock et al. [1] on the reckless practice of machine learning and proprietary prediction models—referred to as “black boxes”—must be acknowledged and possibly extends to all artificial intelligence- and/or mathematical/statistical-related models used in sports medicine and sports sciences settings. The authors conveyed that data-driven methods may pose a threat to effective athlete injury risk prevention and/or performance assessment and modeling. We agree that lack of transparency and external validation raise some ethical concerns (such as data security and privacy, integrity, fairness, algorithmic biases) in smart information system-based predictive analytics that may limit their use in medicine and sports sciences supply chain management (i.e., practitioners’ day-to-day decision-making process). In our view, despite the progress made to connect data science within sports medicine and sports sciences microcosms (e.g. [2,3,4,5]), efforts are still needed to bridge this gap and move towards the convergence of sports and data sciences. Indeed, the current use of machine learning and “black boxes” hinders their understanding, interpretability (i.e., defined as “describing the internals of a system in a way that is understandable to humans” [6]), explainability, and transportability.

Over the past few years, the volume of data has grown at a staggering rate, increasing the need to extract meaning from those numbers in reference to injury risk and performance prediction. However, as stated in Pearl and Mackenzie [7], “causal questions can never be answered from data alone. They require to formulate a model of the process that generates the data, or at least some aspects of that process”. Bullock et al. [1] have warned about “black box” models’ erroneous causal assumptions and called for building models around “a counterfactual, explanatory framework where existing evidence, expert knowledge and clinical reasoning can be used to select predictors considered important both in terms of clinical relevance and to adjust for the effect of confounding factors” [8, 9]. Bearing in mind that correlation implies specific types of association such as monotone trends or clustering, but not systematically causation per se [2, 10], thus excluding any prediction [11], data mining through multivariate models would appear as the first layer in which to search for any interactions between intrinsic and extrinsic risk factors [5] and pose interpretative questions [7]. For instance, low-complexity machine learning models using linear regression, logistic regression, or decision trees are “intrinsically” interpretable due to their (relatively) simple structure: practitioners can quite precisely understand how predictions were made by the model (e.g., by looking at weight attributed to each factor in a linear equation, or decision rules in simple decision trees) [6]. However, to go further in data interpretation, causal models must include contributing factors (i.e., predictors/features/covariates/explanatory variables), confounders, and other pathways that affect outcomes [7]. Actually, it can be assumed that "black boxes" achieve better performance than the aforementioned interpretable models but refer to systems "that do not reveal their internal mechanisms" [12]. This does not facilitate the interpretability of the prediction, thereby legitimating the skepticism about such artificial intelligence models (e.g., [11]).

Answering the question "why" is just as important as predicting, with regard to assisting the decision-making process of practitioners, who try to reduce or prevent injury risk and optimize athletes' performance [13, 14]. Methodologies such as model-agnostics methods [15], which permit extraction of post-hoc information from models ("black boxes" or not), allow us to make a trade-off between transparency, predictive power, and interpretability and help in interpreting complex models. Such methods are not model specific; they can be global or local and provide, for instance, partial dependence or accumulated local effect plots. These may give insight into each feature's influence on outputs, and help build an interpretable model, trained to approximate "black box" predictions.

For example, Bareinboim and Pearl [16] addressed the problem of data fusion in causal inference, providing a complete criterion for deciding on the transportability or not of a result to a new environment, the proviso for use of the criterion being to represent the salient characteristics of the data-generating process with a causal diagram, marked by sites of potential disparity.

In summary, machine-learning methods today are an opportunity to enrich our scope regarding performance or injury-related concerns. Their present use may deserve a red card for reckless practice; all actions that will improve understanding, interpretability, and transportability will tend to demystify the opacity of "black boxes" and increase practitioners' adherence to injury prevention and performance optimization.

References

1. Bullock GS, Hughes T, Arundale AH, Ward P, Collins GS, Kluzek S. Black box prediction methods in sports medicine deserve a red card for reckless practice: a change of tactics is needed to advance athlete care. Sports Med. 2022. https://doi.org/10.1007/s40279-022-01655-6.
2. McCall A, Fanchini M, Coutts AJ. Prediction: the modern-day sport-science and sports-medicine "Quest for the Holy Grail." Int J Sports Physiol Perform. 2017;12(5):704–6.
3. Imbach F, Sutton-Charani N, Montmain J, Candau R, Perrey S. The use of fitness-fatigue models for sport performance modelling:

conceptual issues and contributions from machine-learning. Sports Med Open. 2022;8(1):29.

4. Hulme A, McLean S, Salmon PM, Thompson J, Lane BR, Nielsen RO. Computational methods to model complex systems in sports injury research: agent-based modelling (ABM) and systems dynamics (SD) modelling. Br J Sports Med. 2019;53(24):1507–10.
5. Van Eetvelde H, Mendonca LD, Ley C, Seil R, Tischer T. Machine learning methods in sport injury prediction and prevention: a systematic review. J Exp Orthop. 2021;8(1):27.
6. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA); 2018; 2018. p. 80–9.
7. Pearl J, Mackenzie D. The book of why: The new science of cause and effect. Basic Books, Inc.; 2018.
8. Shmueli G. To explain or to predict? Stat Sci. 2010;25(3):289–310.
9. Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. Chance. 2019;32(1):42–9.
10. Altman N, Krzywinski M. Association, correlation and causation. Nat Methods. 2015;12(10):899–900.
11. Bahr R. Why screening tests to predict injury do not work-and probably never will...: a critical review. Br J Sports Med. 2016;50(13):776–80.
12. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv. 2018;51(5):42.
13. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56.
14. Nassis GP. Leadership in science and medicine: can you see the gap? Sci Med Football. 2017;1(3):195–6.
15. Tulio Ribeiro M, Sameer Singh S, Guestrin C. Model-Agnostic interpretability of machine learning. In: ICML Workshop on Human Interpretability in Machine Learning (WHI). New York; 2016.
16. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. Proc Natl Acad Sci USA. 2016;113(27):7345–52.