International Journal of Sports Physiology and Performance, 2024, 19, 1178-1179
https://doi.org/10.1123/ijspp.2024-0308

Human Kinetics
LETTER TO THE EDITOR

Response to Lolli: Predicting Injuries in Elite Female Football Players With Global-Positioning-System and Multiomics Data

Juan R. Gonzalez[1,2] **and Gil Rodas**[3,4]

[1]ISGlobal, Institute for Global Health, Barcelona, Spain; [2]Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Madrid, Spain; [3]Futbol Club Barcelona Medical Department and Barça Innovation Hub, Barcelona, Spain; [4]Sports and Exercise Medical Unit, Clinic Hospital and Sant Joan de Déu Hospital, Barcelona, Spain

We appreciate the interest and remarks of Dr. Lolli[1] regarding our paper on data modeling of injuries using global-positioning-system and multiomic data.[2]

We agree with the challenges of developing prognostic models from small sample sizes, which can lead to unstable predictions due to data sparsity. Actually, we already acknowledged this limitation in the "Discussion" section of our manuscript. However, we also took steps to mitigate these issues, and we welcome the opportunity to further elaborate on these measures in response to the comments raised in the letter.

First, advancements in integrating genomic, motion tracking, and other technology-based methods for data collection highlight the need for accurate predictive models. However, these data sets often have limited samples due to participant availability, high costs, and logistical challenges of long-term studies in competitive sports.[3] This issue is even more pronounced in our study, which involves professional athletes from a top European football club. Few studies on elite professional teams, especially with female athletes, enhance understanding of their characteristics. Despite complexities in genetics, metabolomics, and training loads, increasing data availability is crucial to benefit from this knowledge and minimize athlete injuries. Such studies are vital for sharing insights with other clubs.

Second, our model fits our cohort well, delivering accurate and actionable predictions for future seasons, as seen in our ongoing work for 2022–23 and 2023–24 with promising results to be published soon. This robustness is evident in the internal validation through cross-validation in our study,[2] which is suitable for small data sets.[4] However, it also highlights the need for each team to calibrate the model with their own data for optimal results (external validation).

Third, our predictive model was developed based on 894 observations derived from distributed lag models created from 14,500 daily observations spanning 2 seasons. Therefore, the effective sample size is 894, rather than being limited to just 24 players. This longitudinal approach yields a more robust data set than what might be inferred from the nominal sample size alone.

On the other hand, we recognize that a limited sample size impacts type S and type M errors, as noted by Lord et al.[5] Dr. Lolli's probabilistic calculations (it is not true as stated in his Table 1) highlight this, but only actual replications in new populations can truly determine the magnitude of the studied effects.[6] This issue, common in sport science[7] and particularly football[8] due to typically low sample sizes, affects our study as well. To address this, our paper includes 95% CIs, indicating the range of the real effect with 95% confidence. For instance, the hazard ratio for polymorphism rs1799750 is estimated at 0.46, ranging from 0.29 to 0.82, indicating a potential 2-fold decreased risk per allele, but possibly only 18% if considering the upper CI limit. This aligns with the theoretical scenarios discussed in the letter.

The interpretation of hazard ratios and the probabilistic index offer complementary insights. Hazard ratios reflect relative risk over time, a common metric in epidemiological studies,[9] while the probabilistic index evaluates the likelihood of one individual experiencing an event before another. As shown in Table 1 of the letter, both methods consistently highlight the statistical significance of the polymorphisms under study, validating their inclusion in the predictive model regardless of the approach used.

In conclusion, this study stands out for integrating traditional global-positioning-system data with multiomics data to advance injury-risk prediction, positioning it as a pioneering effort in the emerging field of sportomics.[10,11] It underscores the feasibility of conducting such complex research and aims to inspire further studies and collaborations to replicate and expand on its findings. The study also underscores the importance of collaboration among sport scientists, clinicians, and coaches. Bridging the gap between personalized medicine and daily clinical practice is crucial for enhancing athlete performance and well-being.

References

1. Lolli L. Comment on González et al: predicting injuries in elite female football players with global-positioning-system and multiomics data. *Int J Sports Physiol Perform*. 2024;19(11). doi:10.1123/ijspp.2024-0246
2. González JR, Cáceres A, Ferrer E, et al. Predicting injuries in elite female football players with global-positioning-system and multiomics data. *Int J Sports Physiol Perform*. 2024;19(7):661–669. doi:10.1123/ijspp.2023-0184
3. Ginevičienė V, Utkus A, Pranckevičienė E, Semenova EA, Hall ECR, Ahmetov II. Perspectives in sports genomics. *Biomedicines*. 2022;10(2):298. doi:10.3390/biomedicines10020298
4. Kuncheva LI, Rodríguez JJ. On feature selection protocols for very low-sample-size data. *Pattern Recognit*. 2018;81:660–673. doi:10.1016/j.patcog.2018.03.012

Rodas https://orcid.org/0000-0002-8060-3914
Gonzalez (juanr.gonzalez@isglobal.org) is corresponding author, https://orcid.org/0000-0003-3267-2146

5. Lord EM, Weir IR, Trinquart L. Design analysis indicates Potential overestimation of treatment effects in randomized controlled trials supporting Food and Drug Administration cancer drug approvals. *J Clin Epidemiol*. 2018;103:1. doi:10.1016/j.jclinepi.2018.06.012
6. Deeks JJ, Higgins JPT, Altman DG.Analysing data and undertaking meta-analyses. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 2019:241–284. Wiley. doi:10.1002/9781119536604.CH10
7. Schweizer G, Furley P. Reproducible research in sport and exercise psychology: the role of sample sizes. *Psychol Sport Exerc*. 2016;23:114–122. doi:10.1016/j.psychsport.2015.11.005
8. Hecksteden A, Kellner R, Donath L. Dealing with small samples in football research. *Sci Med Footb*. 2022;6(3):389–397. doi:10.1080/24733938.2021.1978106
9. Majeed K, Eliyas JK. Design and analysis. In: Eltorai AEM, ed. *Handbook for Designing and Conducting Clinical and Translational Surgery*. 2023:575–589. doi:10.1016/B978-0-323-90300-4.00100-2
10. Muniz-Santos R, Magno-França A, Jurisica I, Cameron LC. From microcosm to macrocosm: the -omics, multiomics, and sportomics approaches in exercise and sports. *OMICS J Integr Biol*. 2023;27:499–518. https://home.liebertpub.com/omi
11. Rodas G, Gonzalez JR. Sportomics: Futbol Club Barcelona's approach to personalized injury prevention. *Apunt Sport Med*. 2024;59:1–2.