

Beyond linear assumptions in pulmonary function analysis: a methodological critique of veteran health outcome models

Yoshiyasu Takefuji*,¹

Faculty of Data Science, Musashino University, Tokyo, Japan

*Corresponding author: Yoshiyasu Takefuji, Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan (takefuji@keio.jp).

To the Editor:

Zeng et al. conducted a comprehensive analysis investigating the clinical utility of pulmonary function testing (PFT) in evaluating long-term outcomes among deployed veterans who maintained preserved spirometry.¹ They employed sophisticated statistical methods, including adjusted mixed-effects regression models and advanced machine learning algorithms, to systematically evaluate how various PFT parameters could predict clinically significant outcomes in this population. Furthermore, they meticulously tracked the longitudinal evolution of PFT patterns in deployed veterans using interval-censored Cox proportional-hazards regression analysis, which allowed for precise temporal assessment of pulmonary function changes despite irregular testing intervals.¹ This methodologically rigorous approach provided valuable insights into the prognostic capabilities of PFT parameters beyond standard spirometric measurements in monitoring respiratory health among veterans with deployment-related exposures.

However, this paper raises critical theoretical and empirical concerns regarding the application of adjusted mixed-effects regression models and Cox proportional-hazards regression. The linear parametric nature of mixed-effects models and the semiparametric structure of Cox regression may be fundamentally misaligned with the inherently nonlinear and nonparametric characteristics of biological respiratory data, leading to erroneous conclusions.^{2–6} While these statistical tools are widely used, Zeng et al. should acknowledge that their reliability hinges on meeting specific underlying assumptions that may not hold in complex biological systems. When linear methods are applied to nonlinear biological relationships, or when parametric/semiparametric approaches are used with data that follows no particular distribution, the resulting interpretations may be significantly distorted.

Supervised models like those employed in their study generate two distinct types of performance metrics: direct prediction accuracy and model-fit metrics. While prediction accuracy can be rigorously validated against ground truth labels from clinical outcomes, other statistical metrics lack corresponding ground truth for validation. These unverifiable assessment metrics—including AIC, BIC, likelihood ratio statistics, intraclass correlation coefficients, Schoenfeld residuals, and proportional hazards

assumptions—are internally calculated constructs without external validation anchors. The reliance on these unanchored metrics can lead to overconfidence in model performance despite potential misspecification of the underlying data structure.

To strengthen their analysis, Zeng et al. should consider implementing complementary nonlinear and nonparametric methodologies to validate their findings, providing a more robust framework that acknowledges the complex, nonlinear nature of pulmonary function and its relationship to clinical outcomes in veterans with deployment-related exposures. This approach advocates for nonlinear nonparametric methods such as Gaussian Process Regression as an alternative to mixed-effects models and Random Survival Forests instead of Cox proportional-hazards regression, potentially revealing patterns that conventional parametric approaches might miss.

Author contributions

Yoshiyasu Takefuji completed this research and wrote this article.

Supplementary material

Supplementary material is available at *Annals of the American Thoracic Society* online.

Conflicts of interest

Please see the ICMJE disclosure forms, which have been provided as [supplementary material](#).

Funding

This research has no fund.

Artificial intelligence disclaimer

No artificial intelligence tools were used in writing this manuscript.

Received: November 9, 2025. Accepted: November 12, 2025

© The Author(s) 2026. Published by Oxford University Press on behalf of the American Thoracic Society. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

References

1. Zeng S, Jani NC, Sotolongo AM, Luo G, Arjomandi M, Falvo MJ. Clinical utility of pulmonary function testing in assessing longitudinal outcomes of deployed veterans with preserved spirometry. *Ann Am Thorac Soc*. 2025;22:1664-1673. <https://doi.org/10.1513/AnnalsATS.202411-1205OC>
2. Parsons S, McCormick EM. Limitations of two time point data for understanding individual differences in longitudinal modeling—what can difference reveal about change? *Dev Cogn Neurosci*. 2024;66:101353. <https://doi.org/10.1016/j.dcn.2024.101353>
3. Hanff AM, Krüger R, McCrum C, Ley C; NCER-PD. Mixed effects models but not *t*-tests or linear regression detect progression of apathy in Parkinson's disease over seven years in a cohort: a comparative analysis. *BMC Med Res Methodol*. 2024;24:183. <https://doi.org/10.1186/s12874-024-02301-7>
4. Quantin C, Asselain B, Moreau T. Le modèle de Cox: limites et extensions [the Cox model: limitations and extensions]. *Rev Epidemiol Sante Publique*. 1990;38:341-356.
5. Jiang N, Wu Y, Li C. Limitations of using COX proportional hazards model in cardiovascular research. *Cardiovasc Diabetol*. 2024;23:219. <https://doi.org/10.1186/s12933-024-02302-2>
6. Xue J, Chen Y, Xue C, et al. Limitations of applying the COX proportional hazards model to glioma studies. *J Transl Med*. 2024;22:1156. <https://doi.org/10.1186/s12967-024-05942-w>