

Letter to the editor: Statistical methodology critique and alternative approaches in H5Nx avian influenza seroprevalence study among French cats

Yoshiyasu Takefuji¹

1. Faculty of Data Science, Musashino University, Tokyo, Japan

Correspondence: Yoshiyasu Takefuji (takefuji@keio.jp)

Citation style for this article:

Takefuji Yoshiyasu. Letter to the editor: Statistical methodology critique and alternative approaches in H5Nx avian influenza seroprevalence study among French cats. *Euro Surveill.* 2025;30(15):pii=2500237. <https://doi.org/10.2807/1560-7917.ES.2025.30.15.2500237>

Article received on 31 Mar 2025 / Accepted on 11 Apr 2025 / Published on 17 Apr 2025

To the editor: Bessi re et al. conducted a seroprevalence study examining cats as sentinels of mammal exposure to H5Nx avian influenza viruses in France from December 2023 to January 2025 [1]. While their analysis using Firth logistic regression suggested that owned, non-hunting cats had 94% lower odds of being H5-seropositive than stray cats (odds ratio = 0.06; 95% confidence interval: 0–0.53; $p = 0.01$), several methodological concerns arise. The study's findings about cats in at-risk d partements showing increased H5 exposure risk ($p = 0.09$) and the disproportionate sampling between at-risk ($n = 23$) and non-risk d partements ($n = 7$) raise questions about the statistical approach.

Most critically, the application of Firth logistic regression in this study appears problematic due to multiple violations of its fundamental assumptions [2–7]. The method requires binary outcome, independent observations, no perfect separation, linearity of logit, absence of multicollinearity and outliers, sufficient sample size, and correct model specification – several of which are violated here. Of particular concern are the imbalanced sampling distribution across departments (23 vs seven) and potential interdependence of observations within geographical clusters, which directly contradict the independence assumption.

Given these limitations, alternative analytical approaches such as Spearman's correlation, Kendall's tau for monotonic patterns [8], or mutual information analysis for complex, nonmonotonic interactions would be more appropriate [9]. These methods are nonlinear nonparametric, incorporating three critical components: data distribution, the statistical relationships between variables, and the validation of statistical significance via p values, making them more suitable for analysing this complex epidemiological dataset.

Ultimately, a deep understanding of these tools and their assumptions is essential for conducting robust and accurate data analyses.

Use of artificial intelligence tools

Not used.

Conflict of interest

The author has no conflict of interest.

References

- Bessi re P, Brun J, Hayes B, Marchand A, Lebouteiller L, Soubies SM, et al. Cats as sentinels of mammal exposure to H5Nx avian influenza viruses: a seroprevalence study, France, December 2023 to January 2025. *Euro Surveill.* 2025;30(12):2500189. <https://doi.org/10.2807/1560-7917.ES.2025.30.12.2500189> PMID: 40156347
- Dey D, Haque MS, Islam MM, Aishi UI, Shammy SS, Mayen MSA, et al. The proper application of logistic regression model in complex survey data: a systematic review. *BMC Med Res Methodol.* 2025;25(1):15. <https://doi.org/10.1186/s12874-024-02454-5> PMID: 39844030
- Pinheiro-Guedes L, Martinho C, O Martins MR. Logistic regression: limitations in the estimation of measures of association with binary health outcomes. *Acta Med Port.* 2024;37(10):697–705. <https://doi.org/10.20344/amp.21435> PMID: 39366365
- Wang T, Tang W, Lin Y, Su W. Semi-supervised inference for nonparametric logistic regression. *Stat Med.* 2023;42(15):2573–89. <https://doi.org/10.1002/sim.9737> PMID: 37161926
- Rifada M, Chamidah N, Ningrum RA. Estimation of nonparametric ordinal logistic regression model using generalized additive models (GAM) method based on local scoring algorithm. *AIP Conf Proc.* 2022;2668(1):070013. <https://doi.org/10.1063/5.0111771>
- Work JW, Ferguson JG, Diamond GA. Limitations of a conventional logistic regression model based on left ventricular ejection fraction in predicting coronary events after myocardial infarction. *Am J Cardiol.* 1989;64(12):702–7. [https://doi.org/10.1016/0002-9149\(89\)90751-0](https://doi.org/10.1016/0002-9149(89)90751-0) PMID: 2801520
- van Maanen L, Katsimpokis D, van Campen AD. Fast and slow errors: Logistic regression to identify patterns in accuracy-response time relationships. *Behav Res Methods.* 2019;51(5):2378–89. <https://doi.org/10.3758/s13428-018-1110-z> PMID: 30187434
- Okoye K, Hosseini S. Correlation tests in R: Pearson Cor, Kendall's Tau, and Spearman's Rho. In: *R Programming*. Singapore: Springer; 2024. https://doi.org/10.1007/978-981-97-3385-9_12
- Gibson JD. Entropy and mutual information. In: *Information Theoretic Principles for Agent Learning*. Cham: Springer; 2024. *Synthesis Lectures on Engineering, Science, and Technology.* https://doi.org/10.1007/978-3-031-65388-9_2

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2025.