Microbial Risk Analysis 31 (2026) 100372

Contents lists available at ScienceDirect

Microbial Risk Analysis

journal homepage: www.elsevier.com/locate/mran

Interpreting genomics-driven microbial criteria: Toward robust and transparent risk models

ARTICLE INFO

Keywords:
Genomics-informed risk assessment
Microbiological criteria
Model biases
Interpretability
Multifaceted approach

Letter to the Editor

Pouzou et al. present an innovative QMRA framework to evaluate the potential impact genomically informed microbiological criteria (MC) could reduce non-typhoidal Salmonella (NTS) illnesses from ground beef (Pouzou et al., 2025). Their study integrates virulence-based serovar grouping with dose-response modeling to analyze scenarios involving test proportion (TP), concentration thresholds (CT), and HV targeting. Testing 100% of production lots at CT = 10 CFU/g was predicted to achieve a 71% reduction in illnesses, while HV-only targeting achieved 46% with minimal diversion. In Pouzou et al., "HV-only" denotes microbiological-criteria scenarios in which test-and-divert actions are applied exclusively to higher-virulence Salmonella serovars, whereas lower-virulence serovars are not subject to diversion. These findings underscore the importance of genomic data in MC design and support strategies aligned with Healthy People 2030 objectives. However, challenges related to interpretability and reproducibility of virulence-based clustering and its integration into dose-response functions warrant further discussion.

Interpretability challenges stem from the use of a tree-based model (e.g., random forest) validated against epidemiological data. Nevertheless, clustering results remain highly sensitive to feature selection, choice of distance metrics, and algorithmic bias—issues extensively discussed in the machine learning literature (Strobl et al., 2007). While the methodological details of the RF-based clustering are documented in Fenske et al. (2023), our concern relates to the methodological choice itself: random forests are not designed to provide mechanistic interpretability, as their internal importance measures arise from ensemble heuristics rather than biologically meaningful structure. These aspects warrant attention because parameter choices can influence group assignments, which in turn affect illness reduction estimates. Epidemiological agreement supports the plausibility of the resulting groups, but it does not establish that the internal feature rankings produced by the RF model constitute mechanistic justification, as predictive agreement and interpretive validity are conceptually distinct.

Several underlying modeling assumptions also bear on the interpretation of the results. Contamination is represented using a Poisson-based homogeneous mixing assumption, which is standard and appropriate in QMRA settings. Diversion is modeled through threshold-based rules, although in practice diversion outcomes may also reflect operational variability that is not explicitly represented in the model. Sampling sensitivity depends on the sample-to-combo mass ratio and diagnostic accuracy, while logistical variability may limit benefits. For example, testing all serovars at CT = 1 would divert ~1 in 19,000 combos, whereas CT = 10 reduces this to ~1 in 91,000, illustrating the trade-off between stringency and feasibility. Incorporating interpretability standards and transparent validation, along with scenario analyses for imperfect diversion or delayed detection, would improve applicability. Critically, external validation remains scarce; replication across independent datasets is needed to strengthen generalizability.

To strengthen interpretability and improve translational potential, we recommend a multifaceted framework. First, unsupervised techniques such as Feature Agglomeration can mitigate multicollinearity and stabilize the feature space (Zhang et al., 2020). Second, model-agnostic screening methods, analogous to Highly Variable Gene Selection, can identify informative taxa without embedding model-specific biases (Xie et al., 2025). Third, feature importance and virulence grouping should be validated using non-parametric statistics (e.g., Spearman's rho, Kendall's tau). A feature should only be considered biologically meaningful if supported by both the explainer and statistical tests with significant p-values. Here, the "explainer" refers to model-agnostic interpretability tools that estimate the contribution of individual features to a model's predictions.

These recommendations are particularly relevant in light of the historical factors that shaped current analytical practice. Random forests rose to prominence partly because the distinction between predictive accuracy and interpretive reliability was often overlooked, leading practitioners to rely heavily on models that performed well in prediction. In contrast, interpretation-oriented classical methods—such as Feature Agglomeration, Highly Variable Gene Selection, or Spearman's rho—serve primarily as feature-selection or characterization tools rather than standalone predictive models, which contributed to their limited adoption despite their strengths in stability and interpretability.

Applying this interpretability framework to HV-only strategies—already shown by Pouzou et al. to achieve a 46% illness reduction with minimal diversion (up to ninefold fewer combos compared to all-serovar

https://doi.org/10.1016/j.mran.2026.100372
Received 24 December 2025; Received in revised form 17 March 2026; Accepted 17 March 2026
Available online 18 March 2026

testing)—would further optimize risk modeling. By combining genomic targeting with transparent validation and robust feature selection, HV-focused MC could become not only effective but also scientifically interpretable and regulatory-ready. Adopting these approaches aligns QMRA with best practices and enhances confidence in genomics-driven interventions (Touw et al., 2013).

In conclusion, their study offers a valuable proof-of-concept for genomics-informed MC in beef safety, but caution is warranted when extrapolating to practice without addressing variability, cost, interpretability, and machine learning bias. Transparency in analytical pipelines—including reporting hyperparameters, cross-validation strategies, sensitivity analyses, and explicit evaluation of bias and feature-selection effects—is essential for reproducibility, regulatory acceptance, and public confidence.

Disclosures

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Souichi Oka: Writing – original draft, Conceptualization. **Kiyo Yoshida:** Investigation. **Yoshiyasu Takefuji:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Fenske, G., Pouzou, J., Pouillot, R., Taylor, D., Costard, S., Zagmutt, F., 2023. The genomic and epidemiological virulence patterns of Salmonella enterica serovars in the United States. PLoS One 18, e0294624. https://doi.org/10.1371/journal.pone.0294624.

Pouzou, J., Costard, S., Pouillot, R., Taylor, D., Zagmutt, F. (2025). Incorporating genomic virulence in a quantitative microbial risk assessment to assess the public health impact of alternative microbiological criteria for Salmonella. Microbial Risk Analysis, 30, 100358. https://doi.org/10.1016/j.mran.2025.100358.

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinform. 8, 25. https://doi.org/10.1186/1471-2105-8-25.

Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., van Hijum, S.A.F.T., 2013. Data mining in the life sciences with random forest: a walk in the park or lost in the jungle. Brief. Bioinform. 14, 315–326. https://doi.org/10.1093/bib/bbs034.

Xie, Y., Jing, Z., Pan, H., Xu, X., Fang, Q., 2025. Redefining the high variable genes by optimized LOESS regression with positive ratio. BMC Bioinform. 26, 104. https://doi.org/10.1186/s12859-025-06112-5.

Zhang, J., Wu, X., Hoi, S.C.H., Zhu, J., 2020. Feature agglomeration networks for single stage face detection. Neurocomputing 380, 180–189. https://doi.org/10.1016/j.neucom.2019.10.087.

Souichi Oka[a,*], Kiyo Yoshida[a], Yoshiyasu Takefuji[b]

[a] *Science Park Corporation, 3-24-9 Iriya-Nishi, Zama-shi, Kanagawa 252-0029, Japan*

[b] *Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan*

[*] Corresponding author at: SciencePark Corporation, 3-24-9 Iriya-Nishi, Zama-shi, Kanagawa 252-0029, Japan.

E-mail addresses: souichi.oka@sciencepark.co.jp (S. Oka), kyoshida@sciencepark.co.jp (K. Yoshida), takefuji@keio.jp (Y. Takefuji).