A038 · Joint Calibration of Multiple Evidence Sources Improves Clinical Variant Classification over Independent Calibration

Presenter: Ross Stewart — Northeastern University

Authors: Ross Stewart, Shantanu Jain, Shawn Fayer, Abbye McEwen, Pankhuri Gupta, Daniel Holmes, Jessica Simon, Jeremy Stone, Sriram Pendyala, A. Felicia Adebanjo, Allyssa Vandi, Raining Wang, Melinda Wheelock, Jordan Opsahl, Madeline Walsh, Aman Shihora, Dustin Maly, Lea M. Starita, Douglas Fowler, Predrag Radivojac

Session: Day 2 · Fri Oct 2, 2026 · 1:15–2:15 PM

High-throughput assays measure variant effects on gene function and are integral to precision medicine. We previously introduced ExCALIBR, an approach to calibrate single functional assays into variant-specific pathogenicity probabilities compatible with ACMG/AMP evidence guidelines. In practice, however, a gene is often characterized by several complementary assays capturing different aspects of function. Calibrating and combining these separately is problematic, as each yields its own evidence and prior, assays are correlated rather than independent, and evidence cannot simply be added or multiplied without violating probability assumptions. Combining these evidence sources requires ad hoc rules (e.g., trusting the strongest signal) that break down when assays disagree. Here we extend ExCALIBR to ExCALIBR-MV, a framework that models multiple assays jointly as a single multidimensional distribution with one shared prior to compute variant-specific probabilities of pathogenicity. We applied ExCALIBR-MV to 40 clinically relevant genes, where concordance with clinical controls (MCC, Matthews correlation coefficient: 0.78 to 0.93) was improved over independent calibration and combination (Figure 1A). This framework naturally extends to computational predictors: jointly calibrating three predictors (REVEL, AlphaMissense, MutPred2) across eight genes improved classification over combining individually calibrated scores (MCC: 0.85 to 0.93; Figure 1B). Because functional and computational evidence are often correlated, we jointly calibrated both evidence types across these same genes, resulting in improved evidence assignment (MCC: 0.88 to 0.95; Figure 1C). These findings support joint calibration of multiple evidence sources as a path toward more reliable clinical variant classification.

Joint Calibration of Multiple Evidence Sources Improves Clinical Variant Classification over Independent Calibration

Ross Stewart[1,*], Shantanu Jain[1], Shawn Fayer[2], Abbye McEwen[2], Pankhuri Gupta[2], Daniel Holmes[2], Jessica Simon[2], Jeremy Stone[2], Sriram Pendyala[2], A. Felicia Adebanjo[2], Allyssa Vandi[2], Raining Wang[2], Melinda Wheelock[2], Jordan Opsahl[2], Madeline Walsh[2], Aman Shihora[2], Dustin Maly[2], Lea M. Starita[2], Douglas Fowler[2], and Predrag Radivojac[1]

[1]*Khoury College of Computer Sciences, Northeastern University, Boston, MA 02115, USA*
[2]*Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA*
[*]*Presenting author*

High-throughput assays measure variant effects on gene function and are integral to precision medicine. We previously introduced ExCALIBR [1], an approach to calibrate single functional assays into variant-specific pathogenicity probabilities compatible with ACMG/AMP evidence guidelines [2, 3]. In practice, however, a gene is often characterized by several complementary assays capturing different aspects of function. Calibrating and combining these separately is problematic, as each yields its own evidence and prior, assays are correlated rather than independent, and evidence cannot simply be added or multiplied without violating probability assumptions. Combining these evidence sources requires ad hoc rules (e.g., trusting the strongest signal) that break down when assays disagree [4]. Here we extend ExCALIBR to ExCALIBR-MV, a framework that models multiple assays jointly as a single multidimensional distribution with one shared prior to compute variant-specific probabilities of pathogenicity. We applied ExCALIBR-MV to 40 clinically relevant genes, where concordance with clinical controls (MCC, Matthews correlation coefficient: 0.78 to 0.93) was improved over combined, independently calibrated evidence (Figure 1A). This framework naturally extends to computational predictors: jointly calibrating three predictors (REVEL, AlphaMissense, MutPred2) across eight genes improved classification over combining individually calibrated predictors (MCC: 0.85 to 0.93; Figure 1B). Because functional and computational evidence are often correlated, we jointly calibrated both evidence types across these same genes, resulting in improved evidence assignment (MCC: 0.88 to 0.95; Figure 1C). These findings support joint calibration of multiple evidence sources as a path toward more reliable clinical variant classification.

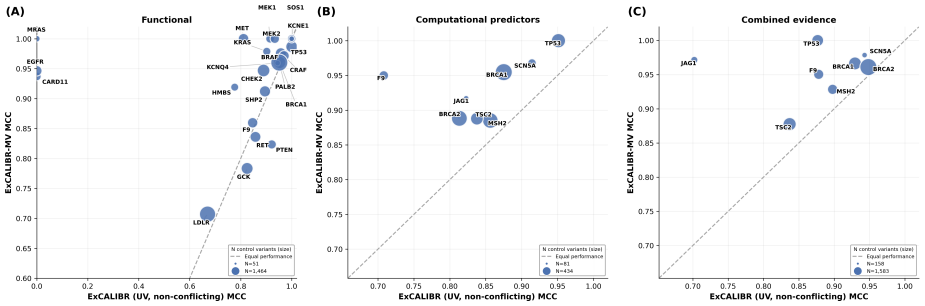

Figure 1: Comparison of ExCALIBR-MV evidence versus ExCALIBR aggregated single-source evidence, for (A) functional assays, (B) variant effect predictors, and (C) combined functional and predictive evidence. The dashed lines denote equal performance.

References

[1] D. Zeiberg et al., "Gene-based calibration of high-throughput functional assays for clinical variant classification," *bioRxiv 2025.04.29.651326*, 2025.
[2] S. Richards et al., "Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology," *Genetics in Medicine*, vol. 17, no. 5, pp. 405–424, 2015.
[3] S. V. Tavtigian et al., "Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework," *Genetics in Medicine*, vol. 20, no. 9, pp. 1054–1060, 2018.
[4] M. Tejura et al., "A scalable approach to resolving variants of uncertain significance," *bioRxiv 2026.02.14.705848*, 2026.