American Journal of Preventive Cardiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

American Journal of Preventive Cardiology

journal homepage: www.journals.elsevier.com/american-journal-of-preventive-cardiology

Short Report

Stability of integrated polygenic and clinical coronary artery disease risk prediction

Dariusz Ratman [a], Robert Maier [a], Akash Kumar [a], Matthew Rabinowitz [a,b], Kate Im [a,*], Akl C Fahed [c,d,**]

[a] MyOme Inc, Menlo Park, CA, USA
[b] Natera Inc., San Carlos, CA, USA
[c] Heart and Vascular Institute, Mass General Brigham, Boston, MA, USA
[d] Broad Institute of MIT and Harvard, Cambridge, MA, USA

ARTICLE INFO

Keywords:
Coronary artery disease
Polygenic risk score
Pooled Cohort Equations
Preventive genomics
Genomic medicine

1. Introduction

Polygenic risk scores (PRS) for coronary artery disease (CAD) are increasingly being deployed to inform preventive cardiovascular care [1–3]. Although PRS show robust population-level associations and are entering clinical workflows [3,4], recent work highlights a key implementation challenge: substantial variability in individual-level risk classification across different PRS models, despite similar discrimination and calibration [5,6].

In a systematic evaluation, cataloged CAD PRSs with comparable population level performance produced discordant individual-level risk estimates, with poor agreement in identifying individuals at the extremes of risk [5]. Thus, stand-alone PRS percentiles may lead to inconsistent high-risk classification.

An ensemble framework has recently been proposed to mitigate this instability and improve classification consistency [6]. However, most PRS are not used in isolation in clinical practice. Current preventive cardiology guidelines rely on multivariable risk equations—such as the Pooled Cohort Equations (PCE)—to estimate 10-year atherosclerotic cardiovascular disease (ASCVD) risk, and PRS are typically conceptualized as risk enhancers or adjunctive predictors rather than stand-alone tools [7,8].

Whether integration of CAD PRS with established clinical risk models improves the stability of individual-level risk prediction has not been systematically evaluated. In this study, we assess the variability of individual risk estimates derived from 14 published CAD PRS and their ensembles, and examine whether combining PRS with the PCE reduces prediction variability and improves agreement in high-risk classification.

2. Methods

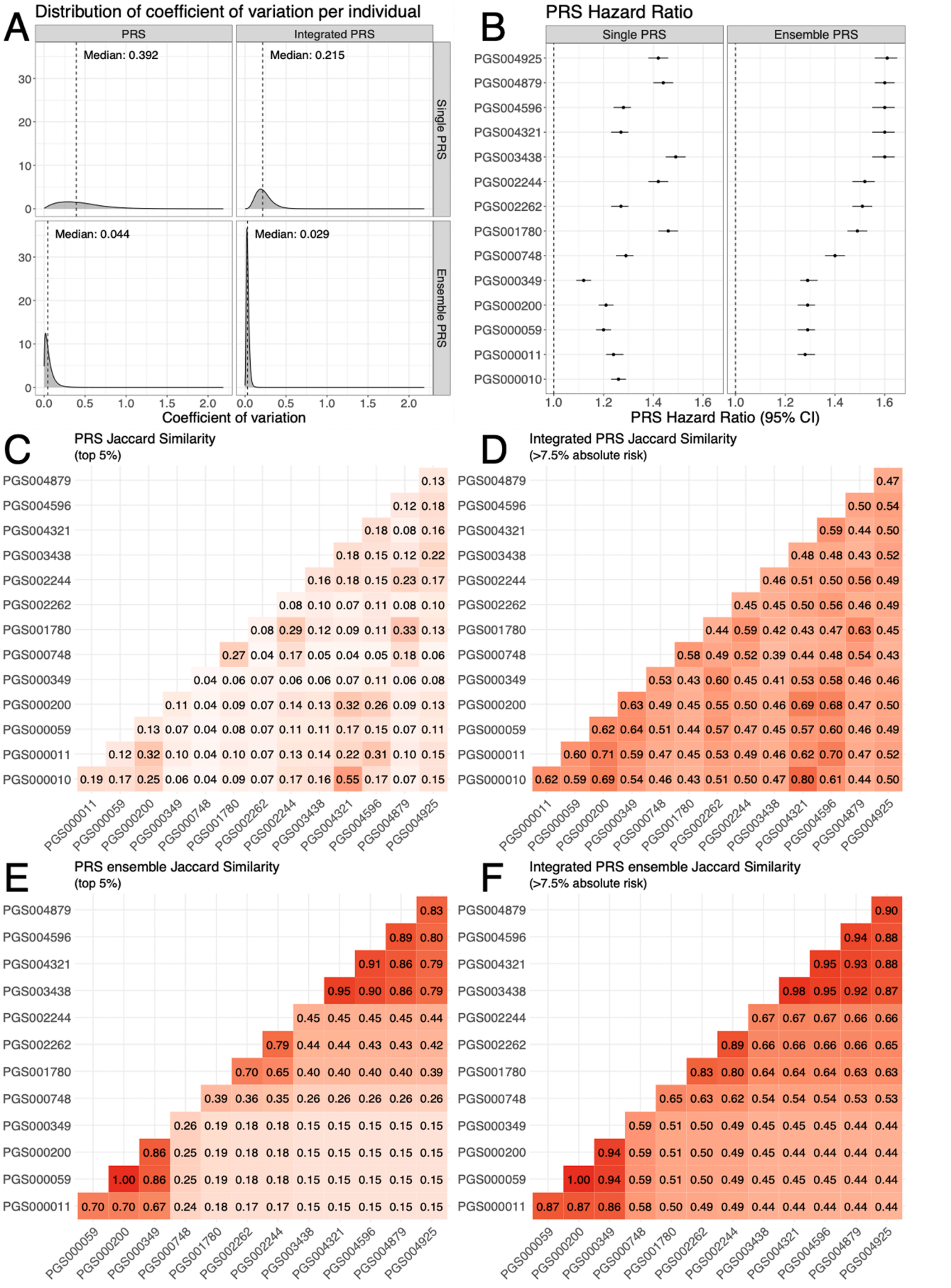
We utilized longitudinal UK Biobank (UKB) data to assess the stability of polygenic risk score (PRS)–only models compared with PRS models integrated with clinical risk factors (Supplementary Figure 1). The analysis included 14 coronary artery disease (CAD) PRSs from the PGS Catalog, each developed independently of UK Biobank GWAS data, as well as ensemble PRS models combining all eligible PGS catalog scores available up to a specific PRS release date [9].

CAD cases were defined using myocardial infarction, coronary revascularization, or death from CAD. Outcomes were censored at 10 years to align with ASCVD-PCE predictions.

PRSs were scored using PLINK and standardized to account for ancestry-specific mean and variance as previously described [10].

* Corresponding author at: MyOme, Inc, 1505 Adams Drive, Suite B1, 940025 Menlo Park, California, USA.
** Corresponding author at: Cardiovascular Research Center, Massachusetts General Hospital, 185 Cambridge Street, CPZN 3.128, Boston, MA 02114, USA.
E-mail addresses: kate@myome.com (K. Im), afahed@mgh.harvard.edu (A.C. Fahed).

https://doi.org/10.1016/j.ajpc.2026.101508
Received 12 February 2026; Received in revised form 24 February 2026; Accepted 25 February 2026
Available online 28 February 2026



(caption on next page)

Fig. 1. Performance and stability of integrated and stand-alone PRS models.
(A) Distribution of individual-level coefficient of variation (standard deviation of individual level predictions divided by their mean) for CAD PRS (left) vs. CAD PRS integrated with ASCVD-PCE (right). The top row shows variation across the five best-performing PGS Catalog models, while the bottom row shows variation in ensemble scores that combine all PGS Catalog models available from the five most recent publication dates. Ensemble CAD PRS models integrated with ASCVD-PCE (bottom right panel) had the least coefficient of variation. **(B)** PCE-adjusted Hazard Ratio (per standard deviation) of single PRSs (left) or ensemble scores (right). Models are ordered based on PGS catalog release date and ensemble includes all models released up to a given PRS release date. **(C)** Binary classification concordance between single PRSs and (**D**) their corresponding integrated PRS with ASCVD-PCE models. (**E**) binary classification concordance between PRS ensembles and (**F**) their corresponding integrated PRS with ASCVD-PCE models. PRS models are ordered based on PGS catalog release date (most recent in the top right of each heatmap) and ensemble includes all models released up to a given PGS release date.

Ensemble PRS weights were optimized using ElasticNet logistic regression. Ensemble training was performed in 13,852 CAD cases at enrollment matched 1:1 with controls by age and sex. For integrated model development and testing, separate subsets of CAD-free participants at baseline were used (n=97,844 for development; n=195,688 for testing), restricted to individuals eligible for ASCVD-PCE risk assessment (Supplementary Figure 1 and Supplementary Table 1).

Integrated models were constructed using Cox proportional hazards regression by combining the ASCVD-PCE linear predictor with standardized PRS additively, with 10-year CAD incidence as the outcome. PCE-adjusted hazard ratios per standard deviation were calculated. Prediction variability was evaluated by transforming PRS to risk percentiles and by calculating absolute 10-year risk for integrated models. The coefficient of variation (CV) quantified dispersion. High-risk agreement was assessed using the Jaccard index, defining high-risk as top 5% of PRS or ≥7.5% 10-year ASCVD risk (~5.8% of the cohort for the integrated ensemble).

3. Results

Across 14 published CAD PRSs, substantial variability was observed in individual-level risk estimates when PRS were used alone. The median coefficient of variation (CV) of individual-level predictions across the top 5 PRSs was 0.392 (Interquartile range [IQR] 0.234-0.586) (Fig. 1A). Integration of each PRS with the ASCVD PCE substantially reduced variability ($p < 0.001$), lowering the median CV to 0.215 (IQR 0.159-0.282).

Combining PRSs into ensemble models further stabilized predictions. Among the top 5 PRS ensembles, the median CV decreased to 0.044 (IQR 0.021-0.081), and to 0.029 (IQR 0.020-0.042) when integrated with PCE (Fig. 1A). Iterative ensemble construction improved population-level performance while reducing dispersion (Fig. 1B). The HR per SD of the highest performing PRS (PGS004925) increased from 1.42 (95% CI 1.38-1.46) for individual PRS to 1.61 (95% 1.56-1.65) in the ensemble framework.

We next evaluated high-risk classification. Using a top 5% threshold for stand-alone PRS percentiles, the median Jaccard index across single PRSs was 0.113 (IQR 0.074-0.165), indicating poor concordance (Fig. 1C). When integrated with PCE and defining high-risk as ≥7.5% 10-year risk (~5.8% of the cohort for the integrated ensemble), the median Jaccard index increased ($p < 0.001$) to 0.497 (IQR 0.457-0.572).

Ensemble PRSs alone improved agreement compared with individual PRSs (median Jaccard index 0.262 [IQR 0.157-0.455]), but stability was highest when ensemble PRSs were integrated with PCE (median Jaccard index 0.59 [IQR 0.456-0.670]) (Fig. 1D).

4. Discussion

In this analysis of 14 published CAD PRSs, we demonstrate that individual-level instability observed with stand-alone PRS is substantially improved when genetic risk is incorporated into a multivariable clinical framework. Single PRSs showed marked dispersion (median CV 0.392) and poor high-risk agreement (median Jaccard 0.113), whereas integration with the ASCVD PCE reduced variability (median CV 0.215) and improved concordance (median Jaccard 0.497). Ensemble approaches further stabilized predictions, with the highest agreement observed when ensemble scores were integrated with PCE (median CV 0.027; median Jaccard 0.59), and near-perfect concordance (0.86–0.97) across successive integrated ensemble iterations.

These findings build on prior work demonstrating substantial discordance in high-risk classification across published PRSs despite similar population-level discrimination [5,6]. Consistent with extreme value theory, correlated models may diverge in distribution tails, producing instability when percentile cutoffs define "high genetic risk" [6]. We show this instability persists for stand-alone PRS but is attenuated when genetic risk is integrated with clinical factors.

Clinically, these findings are most relevant near guideline-relevant therapeutic thresholds, [8] where adoption of newer PRS models could otherwise shift statin eligibility or preventive strategies. Embedding PRS with multivariable model such as PCE dampens variability and yields more stable, interpretable risk estimates.

Finally, ensemble frameworks provide a pragmatic strategy for incorporating newly published PRSs without substantial reclassification. As GWAS sample sizes expand and PRS methods evolve, a dynamic integrative approach may enable updating genetic risk estimates while preserving stability in clinical decision-making.

5. Conclusion

Stand-alone CAD PRSs demonstrate substantial instability in individual-level risk classification despite similar population-level performance. Ensemble modeling and integration with clinical risk prediction reduced variability and improved concordance. These findings support embedding PRS within established risk equations to provide more stable and clinically actionable estimates.

Funding sources

This study was funded by MyOme Inc.

CRediT authorship contribution statement

Dariusz Ratman: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Robert Maier:** Writing – review & editing, Methodology, Formal analysis. **Akash Kumar:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization. **Matthew Rabinowitz:** Writing – review & editing, Methodology, Funding acquisition. **Kate Im:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Akl C Fahed:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Akl Fahed reports financial support was provided by MyOme Inc. Akl Fahed reports a relationship with Goodpath that includes: consulting or advisory and equity or stocks. Akl Fahed reports a relationship with Avigena that includes: equity or stocks. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to acknowledge Dr. Jiayi Sun for her support with data preparation and phenotyping Drs. Tate Tunstall, Lu Yang, and Placede Tshiaba for their input to conceptual discussions related to the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ajpc.2026.101508.

References

[1] Kullo IJ. Clinical use of polygenic risk scores: current status, barriers and future directions. Nat Rev Genet 2025. https://doi.org/10.1038/s41576-025-00900-8. Available at.

[2] Maamari DJ, Brockman DG, Aragam K, Pelletier RC, Folkerts E, Neben CL, Okumura S, Hull LE, Philippakis AA, Natarajan P, Ellinor PT, Ng K, Zhou AY, Khera AV, Fahed AC. Clinical implementation of combined monogenic and polygenic risk disclosure for coronary artery disease. JACC Adv 2022;1:100068.

[3] Lennon NJ, Kottyan LC, Kachulis C, Abul-Husn NS, Arias J, Belbin G, Below JE, Berndt SI, Chung WK, Cimino JJ, Clayton EW, Connolly JJ, Crosslin DR, Dikilitas O, Velez Edwards DR, Feng Q, Fisher M, Freimuth RR, Ge T, , GIANT Consortium, All of Us Research Program, Glessner JT, Gordon AS, Patterson C, Hakonarson H, Harden M, Harr M, Hirschhorn JN, Hoggart C, Hsu L, Irvin MR, Jarvik GP, Karlson EW, Khan A, Khera A, Kiryluk K, Kullo I, Larkin K, Limdi N, Linder JE, Loos RJF, Luo Y, Malolepsza E, Manolio TA, Martin LJ, McCarthy L, McNally EM, Meigs JB, Mersha TB, Mosley JD, et al. Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations. Nat Med 2024;30:480–7.

[4] Patel AP, Wang M, Ruan Y, Koyama S, Clarke SL, Yang X, Tcheandjieu C, Agrawal S, Fahed AC, Ellinor PT, , Genes & Health Research Team, the Million Veteran Program, Tsao PS, Sun YV, Cho K, Wilson PWF, Assimes TL, Heel van DA, Butterworth AS, Aragam KG, Natarajan P, Khera AV. A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. Nat Med 2023;29:1793–803.

[5] Abramowitz SA, Boulier K, Keat K, Cardone KM, Shivakumar M, DePaolo J, Judy R, Bermudez F, Mimouni N, Neylan C, Kim D, Rader DJ, Ritchie MD, Voight BF, Pasaniuc B, Levin MG, Damrauer SM, BioBank Penn Medicine. Evaluating performance and agreement of coronary heart disease polygenic risk scores. JAMA 2025;333:60–70.

[6] Misra A, Truong B, Urbut SM, Sui Y, Fahed AC, Smoller JW, Patel AP, Natarajan P. Instability of high polygenic risk classification and mitigation by integrative scoring. Nat Commun 2025;16:1584.

[7] O'Sullivan JW, Raghavan S, Marquez-Luna C, Luzum JA, Damrauer SM, Ashley EA, O'Donnell CJ, Willer CJ, Natarajan P, American Heart Association Council on Genomic, Precision Medicine, Council on Clinical Cardiology, Council on Arteriosclerosis, Thrombosis and Vascular Biology, Council on Cardiovascular Radiology and Intervention, Council on Lifestyle and Cardiometabolic Health, Council on Peripheral Vascular Disease. Polygenic risk scores for cardiovascular disease: a scientific statement from the American heart association. Circulation 2022;146:e93–118.

[8] Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, Himmelfarb CD, Khera A, Lloyd-Jones D, McEvoy JW, Michos ED, Miedema MD, Muñoz D, Smith Jr SC, Virani SS, Williams Sr KA, Yeboah J, Ziaeian B. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: A report of the American college of cardiology/American heart association task force on clinical practice guidelines. Circulation 2019;140. https://doi.org/10.1161/cir.0000000000000678. Available at.

[9] Lambert SA, Gil L, Jupp S, Ritchie SC, Xu Y, Buniello A, McMahon A, Abraham G, Chapman M, Parkinson H, Danesh J, MacArthur JAL, Inouye M. The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. Nat Genet 2021;53:420–5.

[10] Ratman D, Tshiaba P, Levin M, Sun J, Tunstall T, Maier R, Shah P, Rabinowitz M, Rader DJ, Kumar A, Im K. Polygenic risk scores improve CAD risk prediction in individuals at borderline and intermediate clinical risk. NPJ Cardiovasc Health 2025;2:1–14.