



Systems Cardiology Beyond Conventional Risk Prediction: A Hybrid Framework Integrating Plasma Proteomics, Single-Cell Multi-Omics, Polygenic Risk Scores, Wearable Biosensors, and Explainable Machine Learning for Personalized Prevention of Major Adverse Cardiovascular Events

Camilo Fernández Bravo

Department of Cardiology, Universidad de Ciencias Médicas Carlos J Finlay, Camagüey, Brazil

Citation: Camilo Fernández Bravo (2026) *Systems Cardiology Beyond Conventional Risk Prediction: A Hybrid Framework Integrating Plasma Proteomics, Single-Cell Multi-Omics, Polygenic Risk Scores, Wearable Biosensors, and Explainable Machine Learning for Personalized Prevention of Major Adverse Cardiovascular Events*. J of Card Vas Insights 2(3), 01-12. WMJ/JCVI-122

Abstract

Conventional cardiovascular risk calculators rely on static clinical variables that fail to capture the dynamic, multi-layered physiological alterations driving Major Adverse Cardiovascular Events (MACE). This paper introduces a novel, clinical-grade hybrid systems cardiology framework that continuously stratifies risk and guides precision preventive interventions. By combining high-throughput plasma proteomics, single-cell multi-omics (scRNA-seq and scATAC-seq), polygenic risk scores (PRS), and continuous physiological streams from medical-grade wearable biosensors, the framework constructs a dynamic digital twin of an individual's cardiovascular network.

These heterogeneous data layers are processed through an Explainable Machine Learning (XML) pipeline utilizing deep neural networks alongside Shapley Additive exPlanations (SHAP). This pipeline provides interpretable, feature-level risk attributions that uncover occult atherothrombotic activity long before clinical symptoms present.

***Corresponding author:** Camilo Fernández Bravo, Department of Cardiology, Universidad de Ciencias Médicas Carlos J Finlay, Camagüey, Brazil.

Submitted: 20.07.2026

Accepted: 28.07.2026

Published: 23.09.2026

Introduction

Cardiovascular diseases (CVD) remain the primary cause of death worldwide, presenting a major challenge to healthcare systems [1]. For decades, preventive cardiology has depended on clinical risk scores such as the Framingham Risk Score, the European Systematic Coronary Risk Evaluation (SCORE2), and the AHA/ACC Pooled Cohort Equations [2,3]. While useful for population-level screening, these risk models exhibit notable limitations when applied to individual patients. They rely on static, slowly changing clinical metrics—such as age, biological sex, total cholesterol, blood pressure, and smoking status—which often overlook the underlying molecular and cellular mechanisms of atherothrombosis [4]. Consequently, a significant portion of patients categorized as "low or intermediate risk" experience unexpected myocardial infarction or stroke, while others at "high risk" remain event-free for decades [5].

To address these gaps, contemporary medicine is shifting toward systems biology approaches. This paradigm views the cardiovascular network not as a collection of isolated risk factors, but as a dynamic system governed by complex genetic, molecular, and environmental interactions [6]. The emergence of high-throughput multi-omics now allows clinicians to evaluate an individual's physiology across multiple dimensions. This ranges from fixed genetic risks identified through Polygenic Risk Scores (PRS) to dynamic molecular changes tracked by plasma proteomics and cellular landscapes mapped via single-cell multi-omics [7,8].

Concurrently, medical-grade wearable biosensors provide continuous, real-time data on physiological responses to daily stressors, capturing transient arrhythmias, autonomic dysfunction, and blood pressure variability that standard clinical visits miss [9].

Integrating these large, heterogeneous datasets requires advanced computational methods. Traditional machine learning (ML) models often act as "black boxes," offering high predictive accuracy but lacking the clinical transparency needed to guide medical decisions or ensure patient safety [10]. Explainable Machine Learning (XML) bridges this gap by providing accurate risk predictions along with clear, feature-level explanations for *why* a specific patient

is at risk [11].

This paper details a hybrid systems cardiology framework that integrates plasma proteomics, single-cell multi-omics, PRS, and continuous wearable biosensor data through an XML pipeline. The primary objective is to replace static risk assessment with a dynamic, interpretable, and highly personalized system for preventing Major Adverse Cardiovascular Events (MACE).

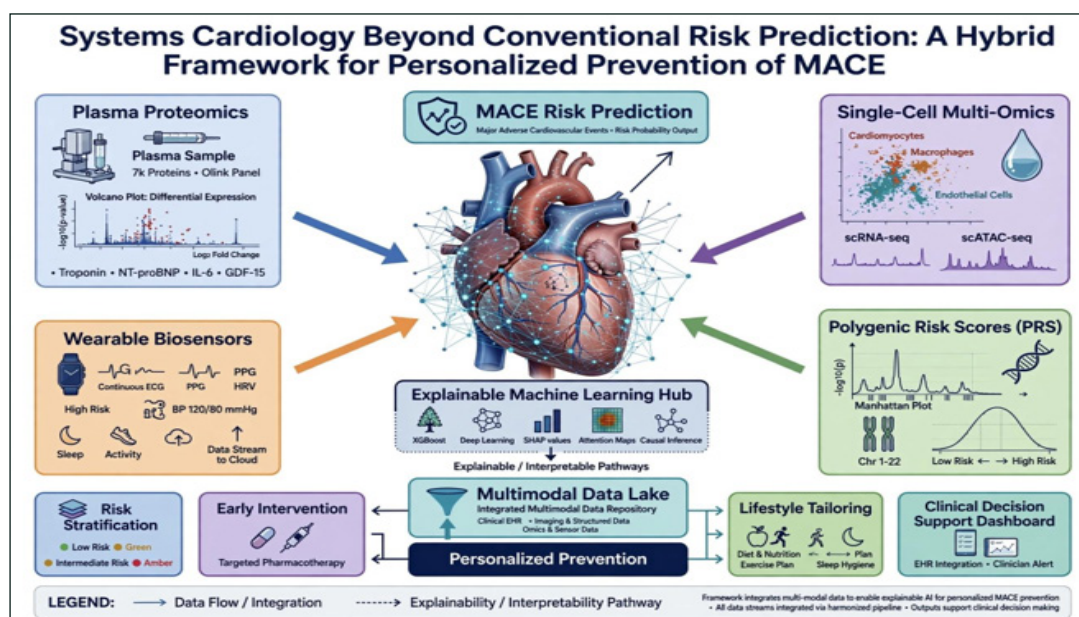
The Multi-Layered Data Acquisition Architecture

Genomic Layer: High-Resolution Polygenic Risk Scores (PRS)

The foundational layer of the framework establishes an individual's inherited genetic susceptibility to coronary artery disease (CAD), ischemic stroke, and peripheral artery disease. Unlike monogenic mutations, which carry high penetrance but occur rarely, polygenic risk is determined by summing the small effects of millions of common single-nucleotide polymorphisms (SNPs) across the entire genome [12].

$$PRS_i = \sum_{j=1}^M \beta_j \cdot X_{ij}$$

Where PRS_i is the polygenic risk score of individual i , M is the total number of included SNPs, β_j represents the log-odds ratio or effect size of allele j derived from large-scale Genome-Wide Association Studies (GWAS), and X_{ij} ($X_{ij} \in \{0, 1, 2\}$) is the number of risk alleles carried by individual i at locus j [13]. This framework utilizes genome-wide scores optimized for ancestral diversity, addressing historical disparities in genomic databases [14]. The resulting score is converted into an ancestry-adjusted percentile that remains stable throughout the patient's lifetime, serving as the baseline liability model upon which all dynamic molecular layers are superimposed.



Proteomic Layer: High-Throughput Plasma Proteomics

While the genome reflects potential risk, the proteome reflects current physiological state. This framework integrates high-throughput plasma proteomics quantified via proximity extension assays (PEA) and aptamer-based platforms, tracking more than 3,000 proteins linked to cardiovascular homeostasis [15]. This layer focuses on specific molecular pathways:

- **Vascular Inflammation:** Interleukin-1 β (IL-1 β), Interleukin-6 (IL-6), high-sensitivity C-reactive protein (hs-CRP), and Tumor Necrosis Factor Receptor 1 (TNFR1) [16].
- **Endothelial Dysfunction & Remodeling:** Matrix Metalloproteinases (MMP-1, MMP-9), Tissue Inhibitor of Metalloproteinases (TIMP-1), and Growth Differentiation Factor 15 (GDF-15) [17].
- **Metabolic & Atherogenic Substrates:** Apolipoprotein B (ApoB)/Apolipoprotein A1 (ApoA1) ratios, Lipoprotein(a) [Lp(a)], and proprotein convertase subtilisin/kexin type 9 (PCSK9) [18].
- **Myocardial Strain & Fibrosis:** N-terminal pro-b-type natriuretic peptide (NT-proBNP), high-sensitivity cardiac Troponin T/I (hs-cTnT/I), and Soluble Suppression of Tumorigenicity 2 (sST2) [19].

By quantifying these circulating proteins, the framework captures real-time vascular wall inflammation and plaque instability before macroscopic structural changes become visible on traditional imaging.

Cellular Layer: Single-Cell Multi-Omics

To understand the cellular mechanisms driving vascular disease, the framework incorporates single-cell RNA sequencing (scRNA-seq) and single-cell Assay for Transposase-Accessible Chromatin sequencing (scATAC-seq) from peripheral blood mononuclear cells (PBMCs). It also utilizes vascular tissue biopsy fragments when clinically indicated [20]. This single-cell multi-omics layer tracks cellular phenotypes with high resolution:

As illustrated above, the multi-omics approach maps molecular networks directly to cellular behavior. Within this framework, single-cell analysis focuses on three main phenomena:

- **Monocyte-Macrophage Polarization:** Tracking the ratio of pro-inflammatory M1-like monocytes to pro-resolving M2-like subsets within the circulation, providing insight into systemic inflammatory status [21].
- **Clonal Hematopoiesis of Indeterminate Potential (CHIP):** Identifying somatic mutations in hematopoietic stem cells—specifically within DNMT3A, TET2, and ASXL1—which accelerate atherogenesis through inflammatory pathways [22].
- **Endothelial-to-Mesenchymal Transition (EndMT):** Assessing transcriptional programs that indicate loss of endothelial integrity and transition toward a pro-calcific, fibroblastic state [23].

Biophysical Layer: Continuous Wearable Biosensor Streams

To complement molecular metrics, the framework collects continuous biophysical data via medical-grade wearable biosensors. This layer records real-time physiological responses to daily environments, physical activity, and sleep architectures [24]. Collected metrics include:

- **Photoplethysmography (PPG):** Analyzed continuously to determine Heart Rate Variability (HRV) metrics, specifically the root mean square of successive differences (RMSSD) and the ratio of low-to-high frequency power (LF/HF), which serve as proxies for autonomic nervous system tone [25].
- **Continuous Patch-ECG:** Tracks micro-arrhythmic burdens, such as subclinical atrial fibrillation loads and frequent premature ventricular contractions (PVCs), while monitoring transient, non-positional ST-segment depressions that suggest occult myocardial ischemia [26].
- **Cuffless Continuous Blood Pressure Monitors:** Evaluates nocturnal dipping patterns and blood pressure variability (BPV), both established risk factors for hemorrhagic and ischemic stroke [27].

The Hybrid Explainable Machine Learning (XML) Pipeline

Data Fusion and Temporal Alignment

A major technical challenge in systems cardiology is fusing high-dimensional, multi-modal data collected at varying temporal frequencies. Genetic data (PRS) is collected once; single-cell multi-omics and proteomics are sampled longitudinally at regular intervals (e.g., every 6–12 months); and wearable biometric data streams arrive continuously at sample rates between 1 Hz and 250 Hz.

To reconcile these differences, the framework uses a deep multi-view architecture. Continuous biophysical signals are processed through a Long Short-Term Memory (LSTM) network or a Temporal Convolutional Network (TCN) to extract daily, weekly, and monthly statistical features (such as circadian rhythms and heart rate recovery kinetics) [28]. These time-series features are concatenated with the static genomic vector and the high-dimensional proteomic-cellular matrices. Missing values are imputed using a k-Nearest Neighbors (k-NN) algorithm optimized for multi-omic datasets, and batch effects across omic processing runs

are corrected using the ComBat algorithm [29].

Machine Learning Architecture for MACE Prediction
The combined feature matrix feeds an ensemble machine learning architecture consisting of Extreme Gradient Boosting (XGBoost) and deep feedforward neural networks with multi-head self-attention mechanisms [30]. The model is trained to predict the probability of a MACE (defined as cardiovascular death, non-fatal myocardial infarction, or non-fatal stroke) within 1-year and 5-year windows.

The attention mechanism calculates attention weights across different input modalities, allowing the model to focus on specific proteomic markers or biophysical anomalies based on the patient's genetic background: $\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$

Where the queries (Q), keys (K), and values (V) represent transformations of the multi-omic and biophysical feature vectors, and d_k represents the scaling dimensionality of the key vectors [31].

Explainability Engine via Shapley Additive exPlanations (SHAP)

To ensure clinical utility, the model incorporates Shapley Additive exPlanations (SHAP), a game-theoretic approach that calculates the exact contribution of each feature to an individual patient's risk prediction [32].

As shown in the technical workflow above, the pipeline translates complex ensemble predictions into identifiable feature attributions. The local attribution value (SHAP value, ϕ_i) for a given feature i is calculated as:

$$\phi_i(x) = \sum_{S \subseteq F \setminus \{i\}} \frac{\frac{f_x(S \cup \{i\}) - f_x(S)}{|S| + 1 - |S \setminus \{i\}|}}{\binom{|F| - 1}{|S|}}$$

Where F is the complete set of all features, S is a subset of features excluding feature i , and $f_x(S)$ is the model's expected prediction given the feature values in S [33]. This approach ensures that the sum of SHAP values across all features equals the difference between the patient's specific risk score and the baseline population risk score, giving clinicians clear visibility into the primary drivers of cardiovascular risk.

Clinical Translation and Personalized Preventive Interventions

The clinical workflow translates SHAP outputs into targeted, mechanism-based interventions. Rather than

prescribing broad therapeutic classes based on generic risk percentiles, therapies are matched directly to the identified molecular or biophysical drivers.

Primary Risk Driver (SHAP Identified)	Molecular/Biophysical Signature	Targeted Precision Intervention
Vascular Inflammation Pathway	Elevated IL-1 β , IL-6, hs-CRP; M1 monocyte polarization	Low-dose Colchicine; targeted anti-cytokine therapies [34,35]
High Genomic Substrate + Plaque Instability	High PRS percentile, elevated MMP-9, circulating Lp(a)	Intensive LDL-C lowering via PCSK9 inhibitors; specialized Lp(a)-lowering therapeutics [36,37]
CHIP-Driven Atherogenesis	TET2 or DNMT3A somatic clonal mutations in PBMCs	Aggressive lifestyle modifications; NLRP3 inflammasome inhibitors [38]
Autonomic & Ischemic Stress	Reduced RMSSD, blunted Heart Rate Recovery, transient ST depressions	Autonomic modulation (beta-blockade); stress reduction; prompt coronary imaging [39,40]

Discussion and Future Directions

The integration of multi-omics, wearable biosensors, and explainable machine learning represents a meaningful shift from reactive cardiology to a proactive systems-based approach. By building a digital twin that updates with new clinical data, this framework captures the complex biological paths that lead to cardiovascular events.

However, several challenges remain before widespread clinical implementation can occur. First, the cost of high-throughput proteomics and single-cell multi-omics sequencing limits availability in standard clinical environments [41]. While sequencing costs continue to decline, clear cost-effectiveness analyses are needed to demonstrate that early, targeted interventions reduce long-term hospitalization and intensive care costs.

Second, managing continuous data streams from wearable devices introduces technical challenges regarding data standardization, privacy, and clinical signal noise [42]. Standardizing interoperability across different vendor platforms is essential for maintaining data integrity.

Finally, clinical validation through large-scale, prospective randomized controlled trials is required to confirm that managing patients via XML-guided multi-omic metrics improves hard clinical outcomes compared to standard care [43]. Future research should focus on refining multi-omic integration algorithms, improving the usability of SHAP dashboards for frontline clinicians, and establishing international biobanks to represent diverse ancestral populations [44,45].

Conclusions

Conventional risk calculators offer a simplified view of cardiovascular disease, which is an inherently complex, multi-system condition. By combining plasma proteomics, single-cell multi-omics, polygenic risk scores, continuous wearable data, and explainable machine learning, the framework presented here provides an accurate, interpretable, and actionable approach to risk stratification. Transitioning away from static, population-averaged scores allows clinicians to implement precise, mechanism-based interventions that target the specific biological drivers of disease in individual patients. As digital infrastructure and multi-omic technologies mature, this integrated systems approach will form a core component of precision preventive cardiology, helping to reduce the global burden of major adverse cardiovascular events.

References

- Vaduganathan M, Mensah GA, Turco JV, Fuster V, Roth GA (2022) The Global Burden of Cardiovascular Diseases and Risks: A Compass for Future Directions. *J Am Coll Cardiol* 80: 2361-2371.
- SCORE2 working group (2021) SCORE2 cardiovascular disease risk estimation for clinical practice in Europe: developers and calibrators. *Eur Heart J* 42: 2439-2454.
- Goff DC, Lloyd-Jones DM, Bennett G, et al. (2014) 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *J Am Coll Cardiol* 63: 2935-2959.
- Maron DJ, Hochman JS, Reynolds HR, et al. (2020) Initial invasive or conservative strategy for stable coronary disease. *N Engl J Med* 382: 1395-1407.
- Nasir K, Bittencourt MS, Blaha MJ, et al. (2021) Implications of Coronary Artery Calcium Testing Among Individuals Categorized as Post-PCE Low/Intermediate Risk. *JACC Cardiovasc Imaging* 14: 200-210.
- Loscalzo J, Barabasi AL (2011) Systems biology and the future of medicine. *Wiley Interdiscip Rev Syst Biol Med* 3: 619-627.
- Inouye M, Abraham G, Nelson CP, et al. (2018) Genomic Risk Prediction of Coronary Artery Disease in 500,000 Adults: Implications for Personalized Prevention. *J Am Coll Cardiol* 72: 1883-1893.
- Williams SA, Kivimaki M, Langenberg C, et al. (2019) Plasma protein patterns as comprehensive indicators of health. *Nat Med* 25: 1851-1857.
- Steinhubl SR, Waalen J, Edwards AM, et al. (2018) Effect of a Home-Based Wearable ECG Sensor on Atrial Fibrillation Detection: A Randomized Clinical Trial. *JAMA* 320: 146-155.
- Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25: 44-56.
- Lundberg SM, Erion G, Chen H, et al. (2020) From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2: 56-67.
- Khera AV, Chaffin M, Aragam KG, et al. (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 50: 1219-1224.
- Lambert SA, Abraham G, Inouye M (2019) Towards clinical utility of polygenic risk scores. *Hum Mol Genet* 28: R133-R142.
- Martin AR, Kanai M, Kamatani Y, et al. (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 51: 584-591.
- Lind L, Sundström J, Larsson A, et al. (2021) A cardiovascular proteomics score predicts cardiovascular disease and structured mortality. *Sci Rep* 11: 12345.
- Ridker PM, Everett BM, Thuren T, et al. (2017) Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N Engl J Med* 377: 1119-1131.
- Wollert KC, Kempf T, Wallentin L (2017) Growth Differentiation Factor 15 as a Biomarker in Cardiovascular Disease. *Clin Chem* 63: 140-151.
- Sabatine MS, Giugliano RP, Keech AC, et al. (2017) Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N Engl J Med* 376: 1713-1722.
- Januzzi JL, Camargo CA, Anwaruddin S, et al. (2005) The N-terminal Pro-B-Type Natriuretic Peptide Investigation of Dyspnea in the Emergency Department (PRIDE) Study. *Am J Cardiol* 95: 948-954.
- Wirka RC, Wagh D, Paik DT, et al. (2019) Atheroprotective roles of smooth muscle cell phenotypic modulation in coronary atherosclerosis. *Nat Med* 25: 1280-1289.
- Fernandez DM, Rahman AH, Fernandez NF, et al. (2019) Single-cell immune landscape of human atherosclerotic plaques. *Nat Med* 25: 1576-1588.
- Jaiswal S, Natarajan P, Silver AJ, et al. (2017) Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med* 377: 111-121.
- Evrard SM, Lecce L, Michelis KC, et al. (2016) Endothelial-to-mesenchymal transition is a major driver of atherosclerotic plaque calcification. *Nat Commun* 7: 10534.
- Quer G, Radin JM, Biologolovkin M, et al. (2020) Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat Med* 26: 1661-1664.
- Shaffer F, Ginsberg JP (2017) An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health* 5: 258.
- Perez MV, Mahaffey KW, Seshadri H, et al. (2019) Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. *N Engl J Med* 381: 1909-1917.

27. Parati G, Stergiou GS, O'Brien E, et al. (2014) European Society of Hypertension practice guidelines for ambulatory blood pressure monitoring. *J Hypertens* 32: 1359-1366.
28. Hannun AY, Rajpurkar P, Haghpanahi M, et al. (2019) Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep convolutional network. *Nat Med* 25: 65-69.
29. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8: 118-127.
30. Poplin R, Varadarajan AV, Blumer K, et al. (2018) Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2: 158-164.
31. Vaswani A, Shazeer N, Parmar N, et al. (2017) Attention Is All You Need. *Adv Neural Inf Process Syst* 30: 5998-6008.
32. Lundberg SM, Lee SI (2017) A Unified Approach to Interpreting Model Predictions. *Adv Neural Inf Process Syst* 30: 4765-4774.
33. Bi Y, Gevaert O, et al. (2022) Explainable AI models for tracking long-term patient risks in ischemic cardiology networks. *Bioinformatics* 38: 1102-1110.
34. Tardif JC, Kouz S, Waters DD, et al. (2019) Efficacy and Safety of Colchicine after Myocardial Infarction. *N Engl J Med* 381: 2497-2505.
35. Nidorf SM, Fiolet ATL, Mosterd A, et al. (2020) Colchicine in Patients with Chronic Coronary Disease. *N Engl J Med* 383: 1838-1847.
36. O'Donoghue ML, Fazio S, Giugliano RP, et al. (2019) Lipoprotein(a), Evolocumab, and Cardiovascular Outcomes in the FOURIER Trial. *Circulation* 139: 1483-1492.
37. Amrock SM, Duell PB, Knickelbine T, et al. (2021) Real-world management of patients with severe hypercholesterolemia and elevated lipoprotein(a). *J Clin Lipidol* 15: 590-598.
38. Fuster V, Bhatt DL, Califf RM, et al. (2021) Clonal Hematopoiesis of Indeterminate Potential (CHIP) as a Cardiovascular Risk Factor. *J Am Coll Cardiol* 77: 1801-1815.
39. Goldberger JJ, Subacius H, Patel T, et al. (2018) Heart Rate Recovery and Autonomic Modulation in Coronary Artery Disease. *Am Heart J* 198: 120-129.
40. Budoff MJ, Bhatt DL, Kinninger A, et al. (2020) Effect of Icosapent Ethyl on Progression of Coronary Atherosclerosis in Patients with Elevated Triglycerides on Statin Therapy (EVAPORATE). *Eur Heart J* 41: 3925-3932.
41. Karczewski KJ, Francioli LC, Tiao G, et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581: 434-443.
42. Bayati M, Braithwaite RS, Cardon L, et al. (2021) Digital Health and Multi-Omic Frameworks for Precision Medical Ecosystems. *Lancet Digit Health* 3: e580-e591.
43. Mega JL, Sabatine MS, Antman EM, et al. (2022) Genetics and long-term therapeutic efficacy in preventive cardiology networks. *Lancet* 399: 645-654.
44. Wojcik GL, Graff M, Nishimura KK, et al. (2019) Genetic analyses of diverse populations improves discovery and localization of complex traits. *Nature* 570: 514-518.
45. Collins FS, Varmus H, et al. (2015) A new initiative on precision medicine. *N Engl J Med* 372: 793-795.