Title: From Data to Wisdom: Harnessing the Power of Multimodal Approach for Personalized Atherosclerotic Cardiovascular Risk Assessment

Authors: Sadeer Al-Kindi, MD and Khurram Nasir, MD, MPH, MSc

Affiliation: Center for Cardiovascular Computational and Precision Health (C3PH), DeBakey Heart and Vascular Center, Houston Methodist, Houston, TX

Address for correspondence
Khurram Nasir, MD, MPH. MSc
Center for Cardiovascular Computational and Precision Health (C3-PH), DeBakey Heart & Vascular Center
Houston Methodist,
6550 Fannin Street, Suite 1801, Houston, Texas 77030.
Phone: (713) 790-3311; E-mail: knasir@houstonmethodist.org; Twitter: @khurramn1

Keywords: Computational cardiology, prediction, prevention

Conflict of Interest: Khurram Nasir: Advisory board of Nova Nordisk, Novartis, Esperion. Speaker Bureau Amgen. Research is partly supported by grants from NIH, Novartis, Esperion, and Jerold B. Katz Academy of Translational Research. Sadeer Al-Kindi has no disclosures.

Funding: No Funding Support for the study.

The influx of data in medical practice has soared, spanning a vast range from electronic health records to diagnostic imaging to genomics. Despite plethora of data at our fingertips, it's astonishing how limited our scope of application remains. Data represents the bottom layer of understanding in the information pyramid, and information, knowledge, and finally intelligence and wisdom build upon data in ascending levels of comprehension and insight. Deliberate analysis and interpretation are thus required to convert raw facts into contextualized knowledge, wisdom, and actionable steps, which can predict future outcomes, guide evidence-based care, and improve population health.¹

Clinical practice has thus far seen only limited utilization of these rich data sources. The capacity for multidimensional thinking is bounded by inherent limitations of human cognition. Consequently, even our most sophisticated risk scores and predictive models seem rudimentary. Outside of medicine, many fields have seen rapid advances thanks to computational analytics. From online recommendations to financial modeling, industries have been revolutionized by repeatedly mining large datasets for patterns and insights that optimize practices. However, medicine has lagged behind in embracing these...
techniques, partly due to the challenges posed by medical data complexity. This variability and disorder make applying traditional statistical methods difficult. There is, however, tremendous potential for computational approaches to similarly transform medical practices.

A glaring manifestation of these current limitations is evident in the prediction of atherosclerotic cardiovascular disease (ASCVD). Almost exclusively, these risk models focus on a narrow set of data points, neglecting the wealth of insights lying at various intersections of patient data. Often, these models are static, reducing their efficacy as they fail to capture the dynamic progression of a patient's health. More sophisticated techniques are needed to uncover multidimensional relationships, interdependencies, and temporal dynamics. As we continually add to our therapeutic arsenal, ensuring that interventions are as precise as possible, treating the right patient at the right time becomes paramount. The urgency for enhanced predictive analysis for ASCVD risk isn't just a matter of technological progress; it's a necessity driven by a confluence of patient needs, economic imperatives, and the overarching goal of improving health outcomes within a resource constraint environment.

In this issue of EHJ Digital Health, Li et al. study of more than 200,000 adults (40-79 years) without ASCVD in China, illuminate the way forward. The authors followed individuals longitudinally and analyzed demographic data, and 25 repeated measures of anthropometric measures and vital signs (e.g. weight, blood pressure), blood markers (lipids, glucose, renal function), medications to predict ASCVD risk (defined as myocardial infarction, stroke, and death due to stroke or coronary artery disease). Summary statistics (such number of measurements, variance and trajectories) were utilized as model inputs. The investigators used two supervised models, eXtreme Gradient boosting (XGBoost) and LASSO regression, to predict the 5-year incidence of ASCVD events. Overall, the ML models achieved good discrimination of ASCVD events (c-statistic of 0.79) and improved discrimination, calibration, reclassification, and clinical utility (measured by decision-curve analysis) as compared with guideline-recommended clinical model based on regression (China-PAR). While the absolute differences in these metrics, as compared with clinical models, are small, they can have large impact when applied to large populations.

As we applaud Li et al efforts, these efforts reflect just the beginning. Looking forward, further refinement of risk prediction will require incorporating diverse multimodal data sources beyond the limited ICD codes, vitals, procedures, labs and medications. Such a narrow viewpoint hinders our ability to truly understand and predict complex diseases like ASCVD. Structured clinical data, though immensely valuable, is just the tip of the iceberg. Beneath this is a depth of unstructured data in the form of clinical notes, reports and imaging findings represent an untapped opportunity when combined with natural language processing. Advanced imaging biomarkers and those derived from biosignals, such as ECG-based risk scores, can provide additional dimensions. Linking social determinants of health and environmental factors with clinical data could also offer unique risk insights (figure). Using proxies like a patient’s home residential address, for example, can yield invaluable data on factors such as social vulnerabilities, area deprivation, air pollution, and other environmental exposures. These elements, often sidelined in traditional analyses, play a profound role in influencing ASCVD risk predictions. Future advancements in risk prediction clear hinge on assimilating diverse data sources. Additionally, longitudinal tracking of risk trajectories over time using timeseries data could unlock new temporal dynamics. With exponential growth in computational power, advanced analytics applied to interconnected, multimodal patient data at scale will enable a new paradigm of preventive medicine with unprecedented personalization and precision.
But “with great power comes greater responsibility”. While confluence of medical data and computational prowess offers untapped potential in reshaping the landscape of CVD risk prediction, it's imperative to also exercise caution and rigor. New techniques might unveil potential risk factors previously uncharted and unprecedented depth in data analysis, they do not negate the need for human expertise. Although machine learning is adept at identifying subtle correlations in datasets, we must remember that correlation isn’t causation. We must strive to understand the foundational reasons for these correlations to ensure their relevance and appropriateness.

In conclusion, the pioneering work of researchers like Li and others has shown us the direction that with determination, innovation, and the right tools, we can navigate these hurdles. However, we would like to also remind that this transformation does not need to be just a technical one; it’s imperative that its deeply human. The future of these efforts has to be about harnessing the power of data and technology to impact lives, alleviate suffering, promoting health, and enhance resource utilization. Every byte of data, algorithm, and predictive model should serve a singular purpose: to enhance the human experience of both the caregiver and the receiver. Adhering to this fundamental principle, while harnessing the full potential of our data resources and integrating them with cutting-edge computational methods, we can set the stage for a new era of medicine – one that's truly personalized, proactive, and precise, but also centered on human values and connections.

References:


Figure: Multimodal Computational Platforms for Precision Prediction of Cardiovascular Risk Across the Knowledge Continuum