A Machine Learning Framework to Identify Ecological Risk Pathways in Cardiovascular Stress: Insights for Health Equity Using Decision Trees and SHAP

Marcia E. I. Uddoh MD (candidate), PhD, MPH, MS, MSW

Black Catholic Health Research Institute | Stress Vitals Institute LLC

INTRODUCTION

Cardiovascular disease (CVD) resulting from chronic stress has been consistently linked to increased morbidity and mortality. Recognized by the World Health Organization as a critical intermediate social determinant of health (Solar, 2013), chronic stress demands targeted attention. Biomarkers play an essential role in the early detection of CVD, which disproportionately affects Black populations. However, stress is frequently assessed through subjective measures, offering limited actionable insights for clinicians and public-health officials aiming for precision in understanding disease progression. As noted by Gaffey et al. (2022), this is especially problematic in cardiovascular health, where many remain hesitant to integrate psychological stressors into formal clinical guidelines. In contrast, biological and physiological markers can provide objective assessments of stress, enabling earlier identification of risk and the development of timely interventions to improve health outcomes. Particularly within Black communities, stress emerges from the dynamic interplay of complex environmental factors operating across multiple levels of the social-ecological model. As highlighted by Golden and Earp (2012), the health-promotion field has often focused narrowly on lifestyle changes while overlooking the broader contextual forces shaping health outcomes. This model offers an effective framework by mapping stressors at the individual, patient–physician, institutional, community, and policy levels, each potentially linked to specific biomarkers indicative of cardiovascular-disease risk patterns. Once identified, these biomarkers can be applied in clinical and public-health settings to enable early detection of disease, whether at the level of an individual patient profile or across broader populations.

METHODS

Data used for this research was provided by the longitudinal study “Midlife in the United States” (MIDUS), managed by the Institute on Aging, University of Wisconsin, and supported by the National Institute on Aging (P01-AG020166). From MIDUS II we analyzed the Milwaukee cohort—an exclusive sample of African American participants with comprehensive biomarker profiles. All data handling adhered to strict de-identification protocols, and all data remained local to our computing environment to ensure privacy and security. Our dataset included a range of biomarkers associated with cardiovascular disease, including serum markers (C-reactive protein [CRP], interleukin-6 [IL-6], and tumor necrosis factor-alpha [TNF-α]). We also incorporated physiological indicators of autonomic function, such as baroreflex sensitivity, heart rate (HR), and heart-rate variability (HRV) metrics relevant to stress regulation and cardiovascular health: root mean square of successive differences (RMSSD), SDRR, and low-frequency (LF) power. Additional health measures, such as body-mass index (BMI), were also included. To construct our ecological framework, we developed a set of indicators across five distinct ecological levels of health. The internal consistency of these ecological indicators was verified using Cronbach’s alpha, with all levels demonstrating acceptable reliability ($\alpha > 0.69$). In our analysis, X represented the full set of biomarkers, while Y corresponded to the ecological levels of health. The dataset was partitioned into training (80 %) and testing (20 %) subsets. We trained and evaluated various decision-tree algorithms to assess how effectively the model could classify validated ecological-level clusters based on patient-specific biomarker profiles.

Scan to Meet slimmAI: Your Conversational HRV Coach

Scan for live SlimmAI demo

RESULTS

Dataset Overview: Our study sample comprised 142 participants, all self-identified as Black adults. We collected a comprehensive set of measurements, including cardiovascular biomarkers—CRP, IL-6, and TNF-α—alongside physiological measures such as HR, HRV, and BMI. In addition to biological data, we captured psychosocial factors relevant to cardiovascular risk, specifically financial stress and experiences of discrimination. At the policy level, insurance coverage through Medicare and Medicaid was included to reflect structural determinants of health. Moving forward, further analysis will target HRV metrics mapped to specific ecological levels to deepen our understanding of autonomic function in stress-related cardiovascular risk.

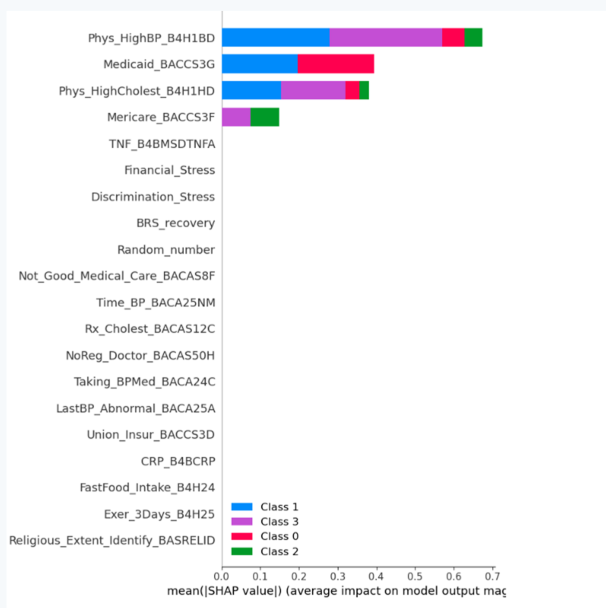
Model Performance: We applied a decision-tree classifier that incorporated variables across ecological levels—biological, physiological, psychosocial, and multilevel attributes. The model achieved an accuracy of 97 % and a macro-averaged F1-score of 0.97 across three classes. The confusion matrix showed near-perfect precision for one class and strong performance for the others (F1-scores of 0.97 and 0.93), with only one misclassification overall. These results demonstrate the model’s effectiveness in handling multiclass ecological classification with minimal error, even across complex ecological combinations.

Key Findings from SHAP Analysis: To enhance model interpretability, we applied SHAP to identify the most influential features across ecological levels. Fast-food intake emerged as the top predictor, followed by physician-recorded high blood pressure, high cholesterol, Medicaid coverage, financial stress, discrimination, baroreflex-sensitivity (BRS) recovery, limited access to quality medical care, timing of last blood-pressure check, and cholesterol prescriptions. Policy-level factors—particularly Medicaid and Medicare—were major contributors to model performance. This aligns with prior findings that Medicaid programs often have limited incentives for early cardiovascular-disease management, as the costs of prevention fall on Medicaid while long-term savings are realized by Medicare (Trogdon et al., 2007). Psychosocial stressors, notably financial stress and discrimination, also showed significant influence on ecological risk levels.

Extended SHAP Insights: HRV metrics did not emerge as dominant contributors in the model and will be examined further in targeted analyses. Our findings suggest that structural policy factors and behavioral determinants carry greater weight in social-ecological models of health, where environmental and systemic structures often outweigh individual-level factors. This aligns with observations by Stokols (1996), who noted that health-promotion efforts frequently lack comprehensive theoretical frameworks, focusing narrowly on individual behavior change while overlooking broader environmental drivers of health. Importantly, our model successfully identified combined ecological patterns, capturing the complex interactions across multiple levels that shape cardiovascular risk. Figure 1 illustrates these interactions, highlighting how key indicators align across ecological levels within the model’s classification. Figure 1. (see supplementary poster) SHAP summary plot highlighting key contributors to ecological-level classification. Fast-food intake and policy-level factors (Medicaid, Medicare) were top drivers of model predictions.

CONCLUSION

We developed a machine-learning model that effectively identifies the BaCE Pathway, providing ecological-level explainability for cardiovascular biomarkers. Our approach captures the complex ecological contexts in which chronic stress contributes to cardiovascular risk through objective biomarker profiles at both individual and population levels. As a next step, we will expand our model to focus on HRV, aiming to better understand autonomic responses to stressors across ecological tiers. Persisting health disparities, particularly among Black populations, highlight the limitations of isolated biomarker analysis that overlooks environmental and systemic factors. By leveraging data-driven visualizations and context-specific biomarker patterns, our model offers actionable insights to advance health equity. This framework empowers clinicians and public-health practitioners to more precisely target interventions across ecological levels. Future work includes external validation and integration into clinical and public-health decision-support systems to enhance precision-prevention strategies. SlimmAI now streams Polar-strap HRV data to Kubios, which Robynn—our conversational LLM—interprets. Pilot tests show Robynn delivers and adapts image-based stress prompts using pre/post-HRV shifts.



ACKNOWLEDGEMENTS

Data used for this research was provided by the longitudinal study titled “Midlife in the United States,” (MIDUS) managed by the Institute on Aging, University of Wisconsin. This research was supported by a grant from the National Institute on Aging (P01-AG020166).