A machine learning approach integrating social determinants of health and allostatic load to predict dysglycemia

PRESENTER

Dante Anthony **Tolentino**, PhD, RN

UCLA Joe C. Wen School of Nursing

Muntasir Masum, PhD

University at Albany State University of New York

Background

- SDOH or conditions of life contribute to poorer health outcomes and are established risk factors for dysglycemia.
- Adverse SDOHs act as chronic stressors, leading to AL, a measure of cumulative physiological "wear and tear" across multiple body systems.
- Elevated AL → increased risk of chronic diseases, including dysglycemia.
- Existing diabetes risk models often overlook this social context and physiological stress.

Purpose

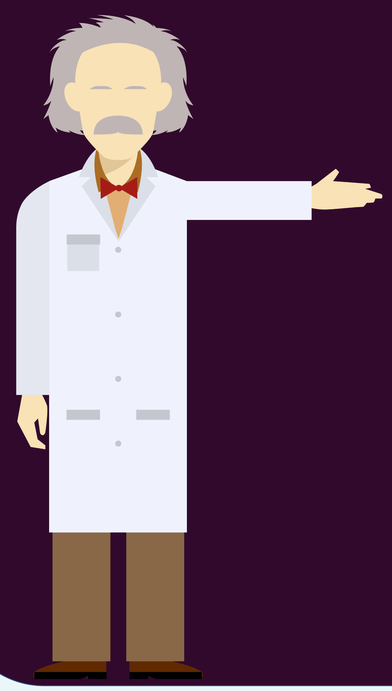
- To develop a machine learning (ML) model that combines social determinants of health (SDOH) with allostatic load (AL) biomarkers to predict dysglycemia (prediabetes and type 2 diabetes).

Methods

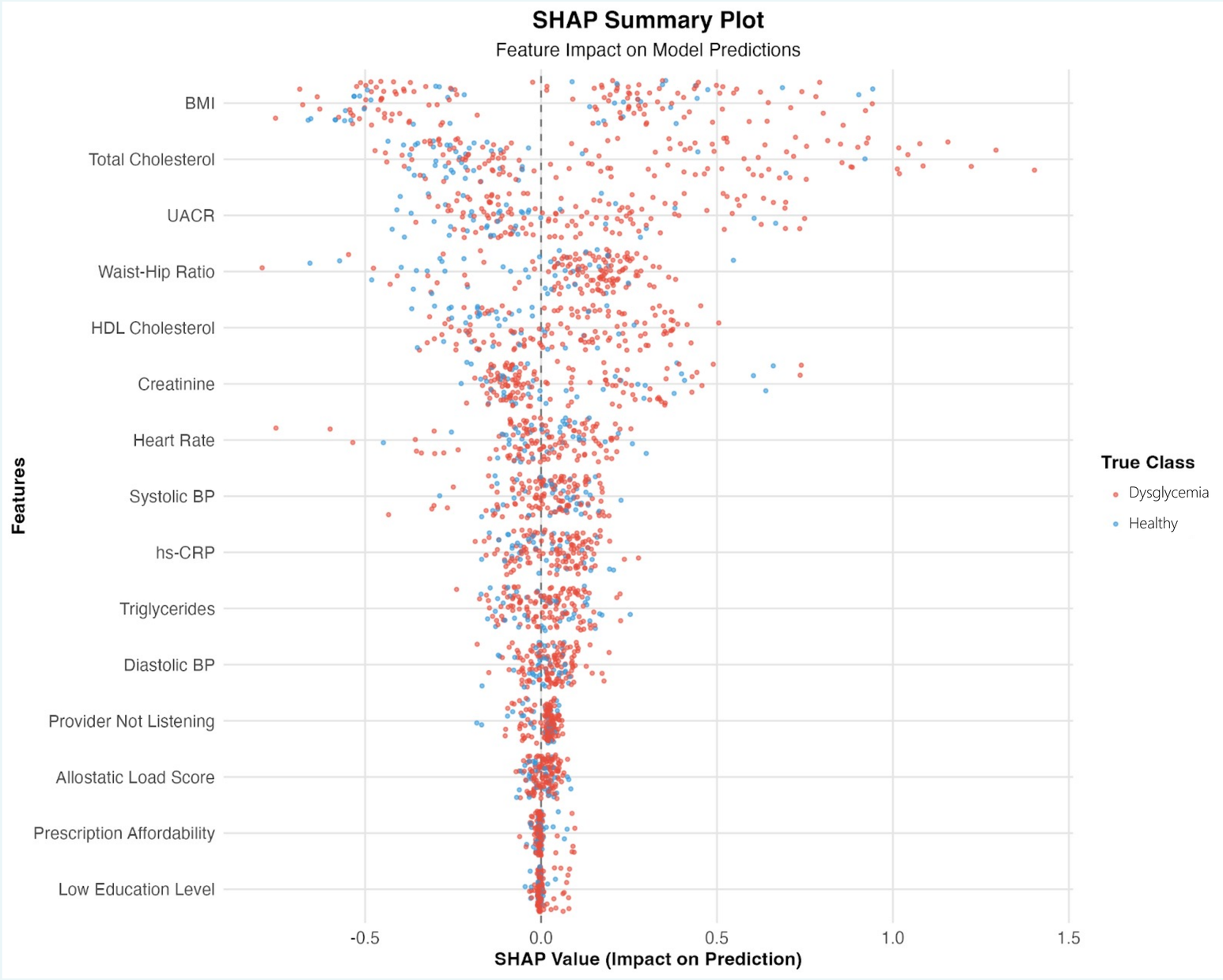
- Supervised ML model (XGBoost) using cross-sectional data from AI-READI (N = 1067, age > 40 years).
- **Outcome**: Binary classification – normoglycemic (n = 372) vs. dysglycemic (n = 695).
- **Features**: 17 features, including 5 SDOH high-risk indicators (e.g., delayed medical care, low education), 11 clinical biomarkers (e.g., BMI, SBP, HDL), and AL score.
  - AL calculation: Count-based approach, summing the number of 11 biomarkers exceeding established clinical thresholds (range 0-8)
- **Model**: Gradient-boosted decision tree classifier (XGBoost) trained on 70% of the data with a held-out test set (15%) for final validation
- **Interpretability**: Model performance assessed using accuracy, precision, recall, F-1 score, ROC-AUC. Feature importance was determined using the SHapley Additive exPlanations (SHAP analysis; see Figure)

Conclusion

- Demonstrated that ML model achieve strong performance in dysglycemia prediction (All model performance scores > 80%)
- This approach supports the value of integrated risk assessment for more holistic diabetes prevention and management strategies.
- Future work should focus on longitudinal and external validation.



The ML model achieved a strong performance in predicting dysglycemia (Accuracy: 80.8%, ROC-AUC: 78.4%), with clinical biomarkers dominating predictive power but AL and SDOH providing meaningful additional value.



This work is supported through the NIH AIM-AHEAD Bridge2AI Program (Award number 1OT2OD032644) and the UCLA Hellman Fellowship.