

# CGM Foundation Model Trained on Largest Healthy Cohort Outperforms Medical SOTA at Diabetes Prediction

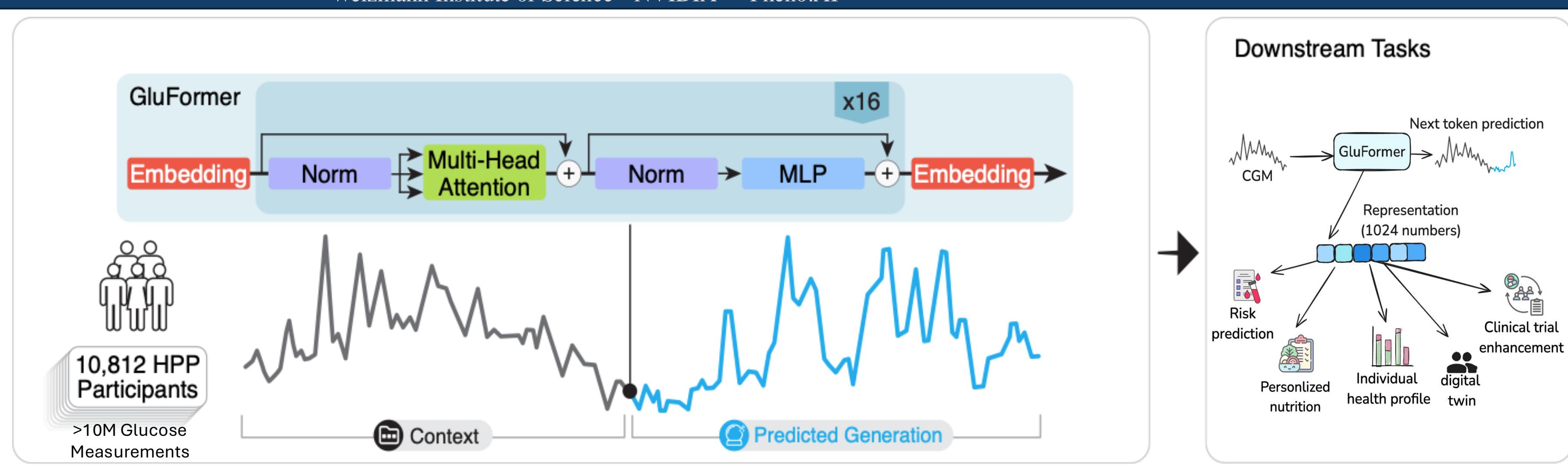


NVIDIA.

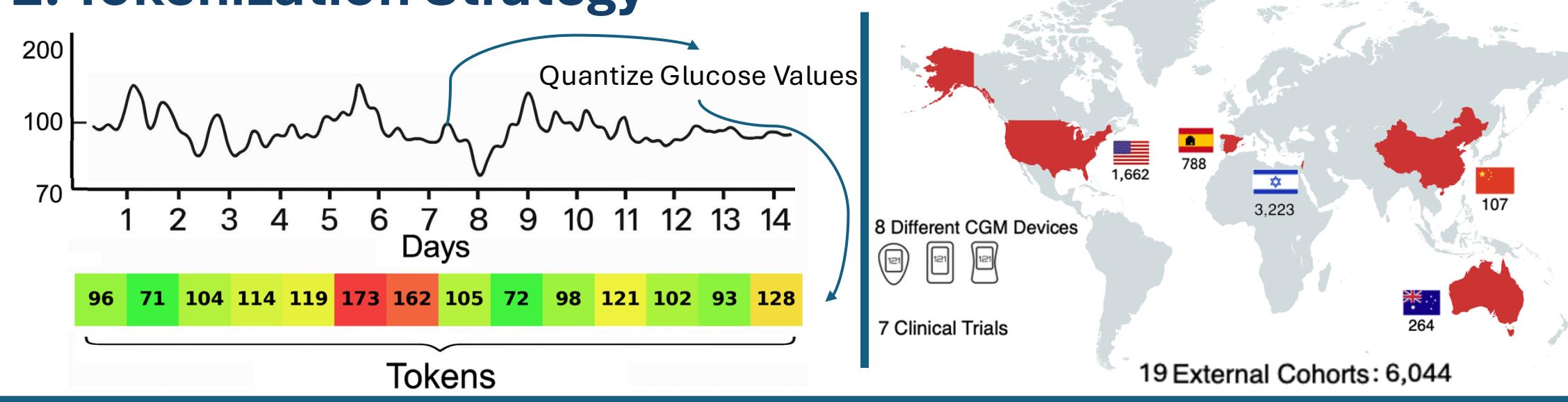
Guy Lutsker<sup>1,2</sup>, Gal Sapid<sup>1,3</sup>, Smadar Shilo<sup>1</sup>, Jordi Merino, Anastasia Godneva<sup>1</sup>, Jerry R Greenfield, Dorit Samocha-Bonet, Raja Dhir, Francisco Gude, Shie Mannor<sup>2</sup>, Eli Meiron<sup>2</sup>, Gal Chechik<sup>2</sup>, Hagai Rossman<sup>1,3,\*</sup> & Eran Segal<sup>1,\*</sup><sup>1</sup> Weizmann Institute of Science <sup>2</sup> NVIDIA <sup>3</sup> Pheno.AI

## 1. Main Idea

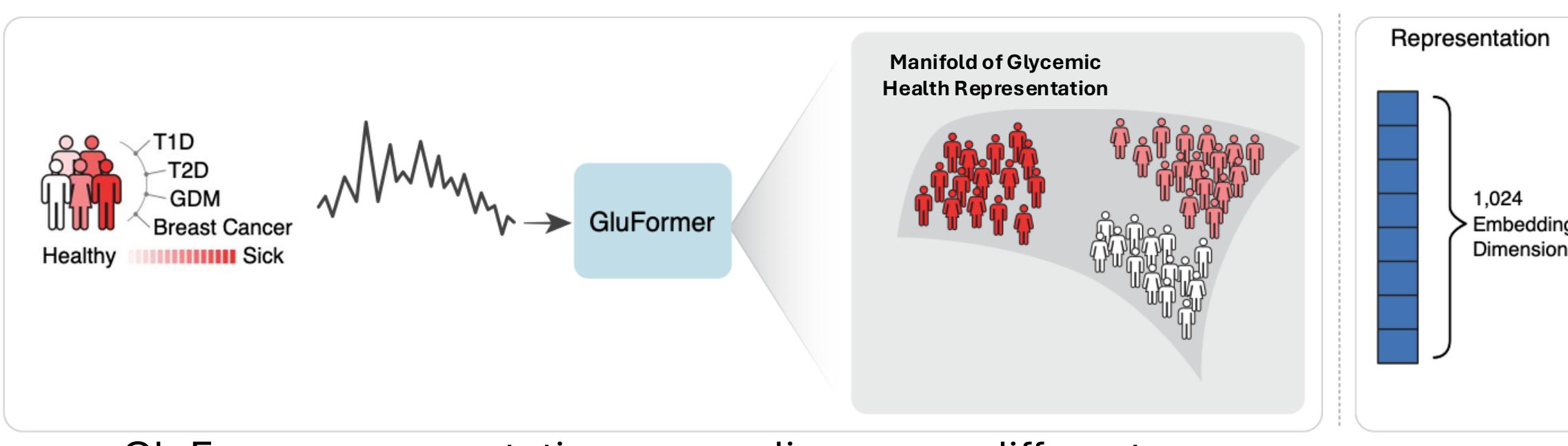
Use a cohort of ~10,000 individuals with continuous glucose monitoring (CGM) to build a generative transformer model using next token prediction.



## 2. Tokenization Strategy



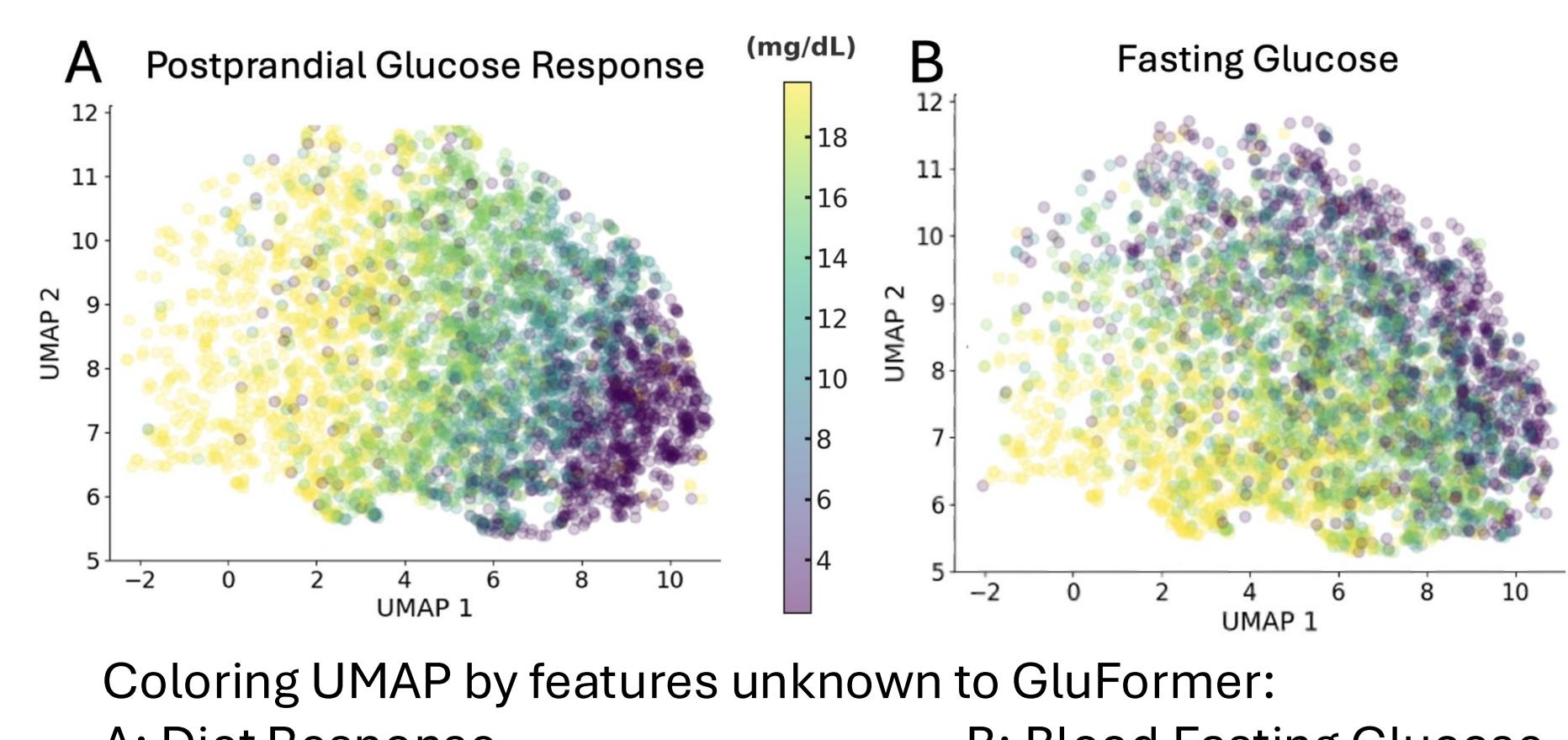
## 3. Creating Generalizable Representations: Validated on OOD Cohorts



GluFormer representations generalize across different axes:

- Diseases.
- CGM devices.
- Populations over the world.

## 4. Representation Space Captures External Phenotypes

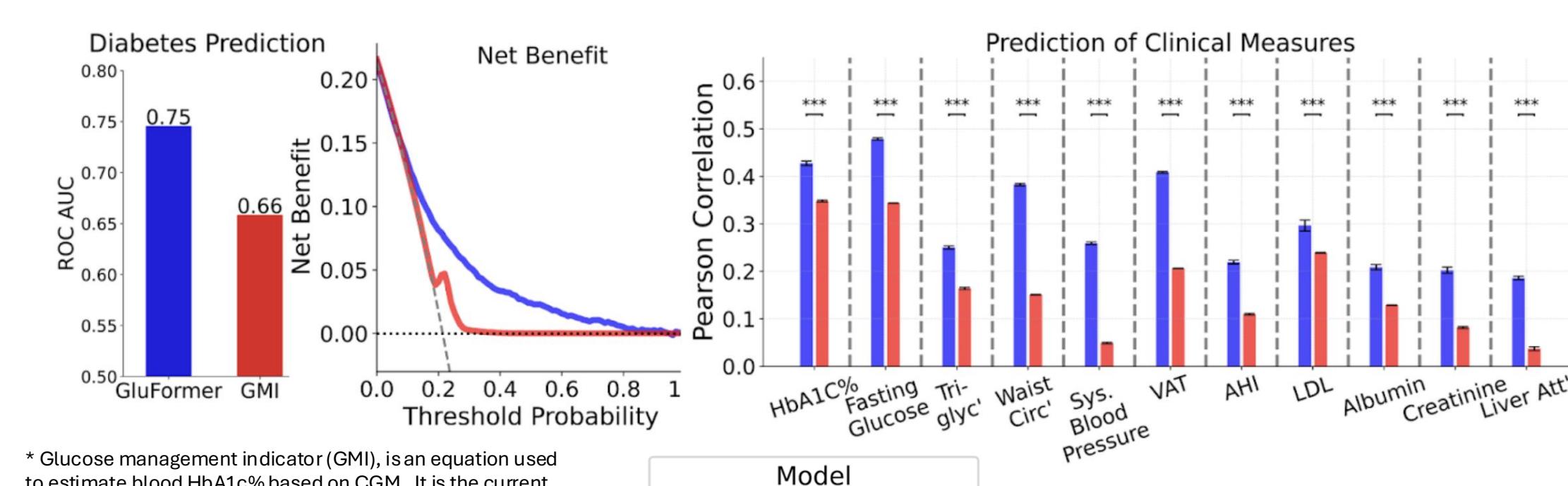


Coloring UMAP by features unknown to GluFormer:

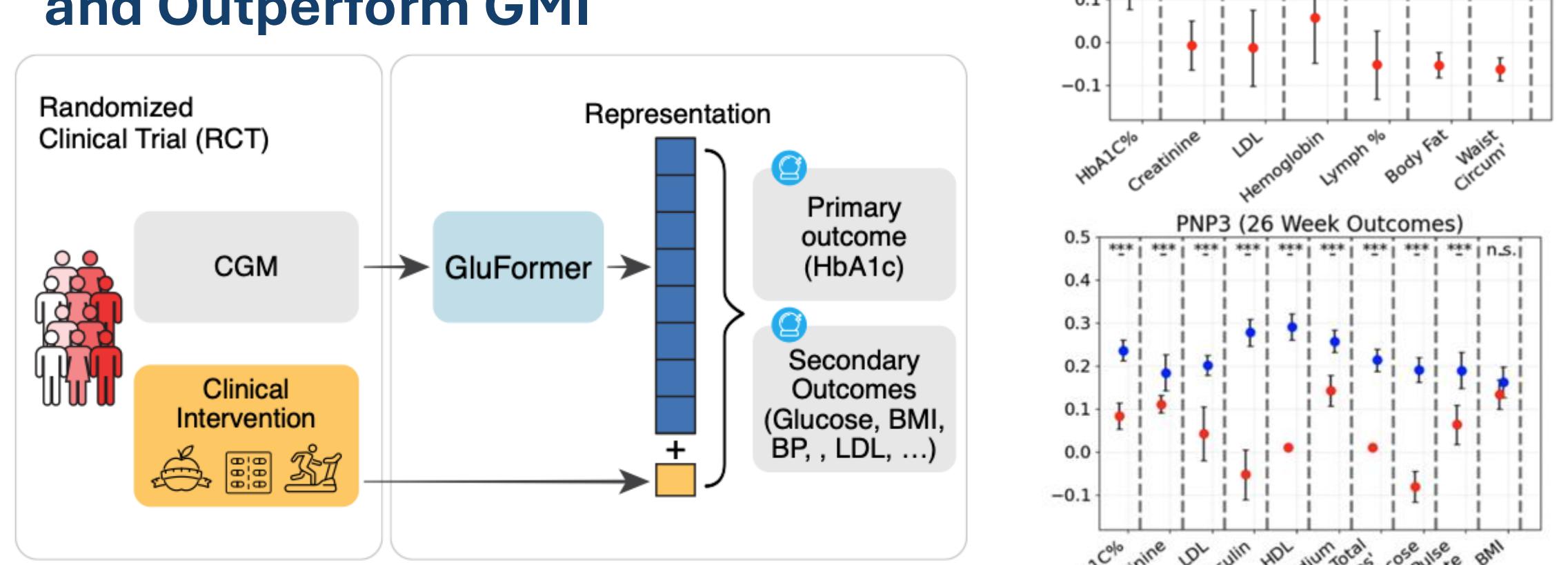
A: Diet Response

B: Blood Fasting Glucose

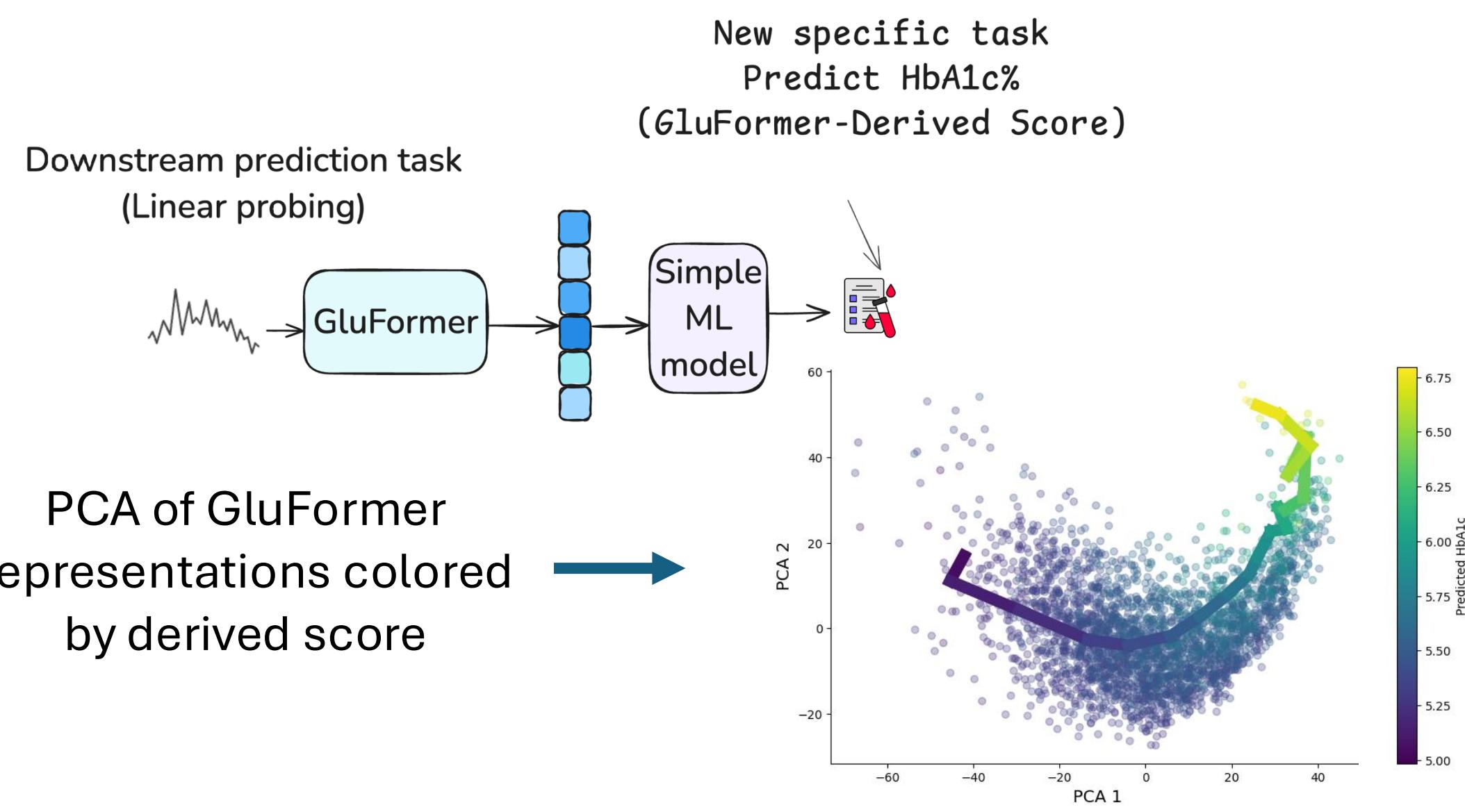
## 5. Representations are Predictive of General Health, and Outperform Gold-Standard CGM Summary Tool - GMI



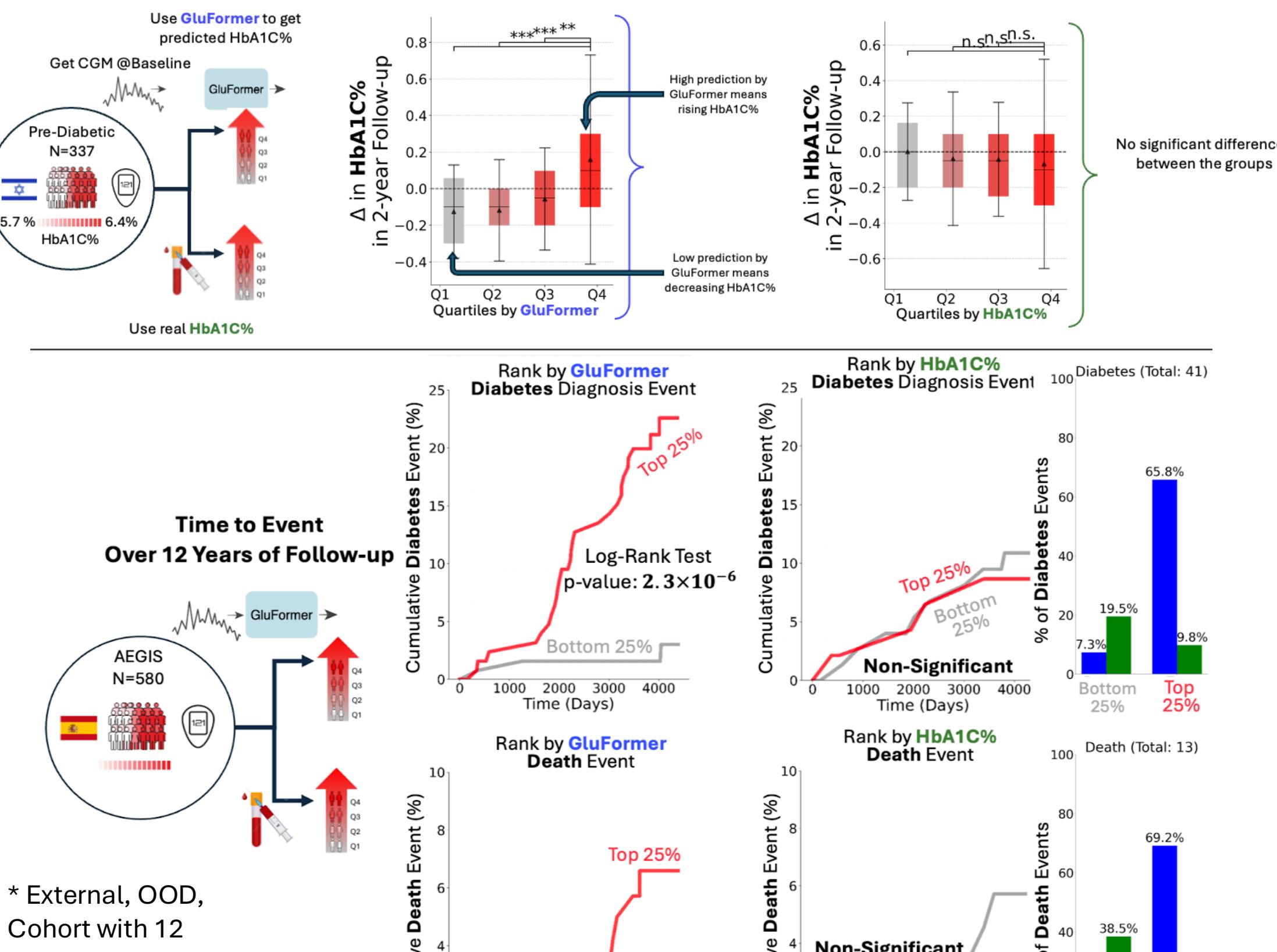
## 6. Representations are Predictive of Randomized Clinical Trial Results, and Outperform GMI



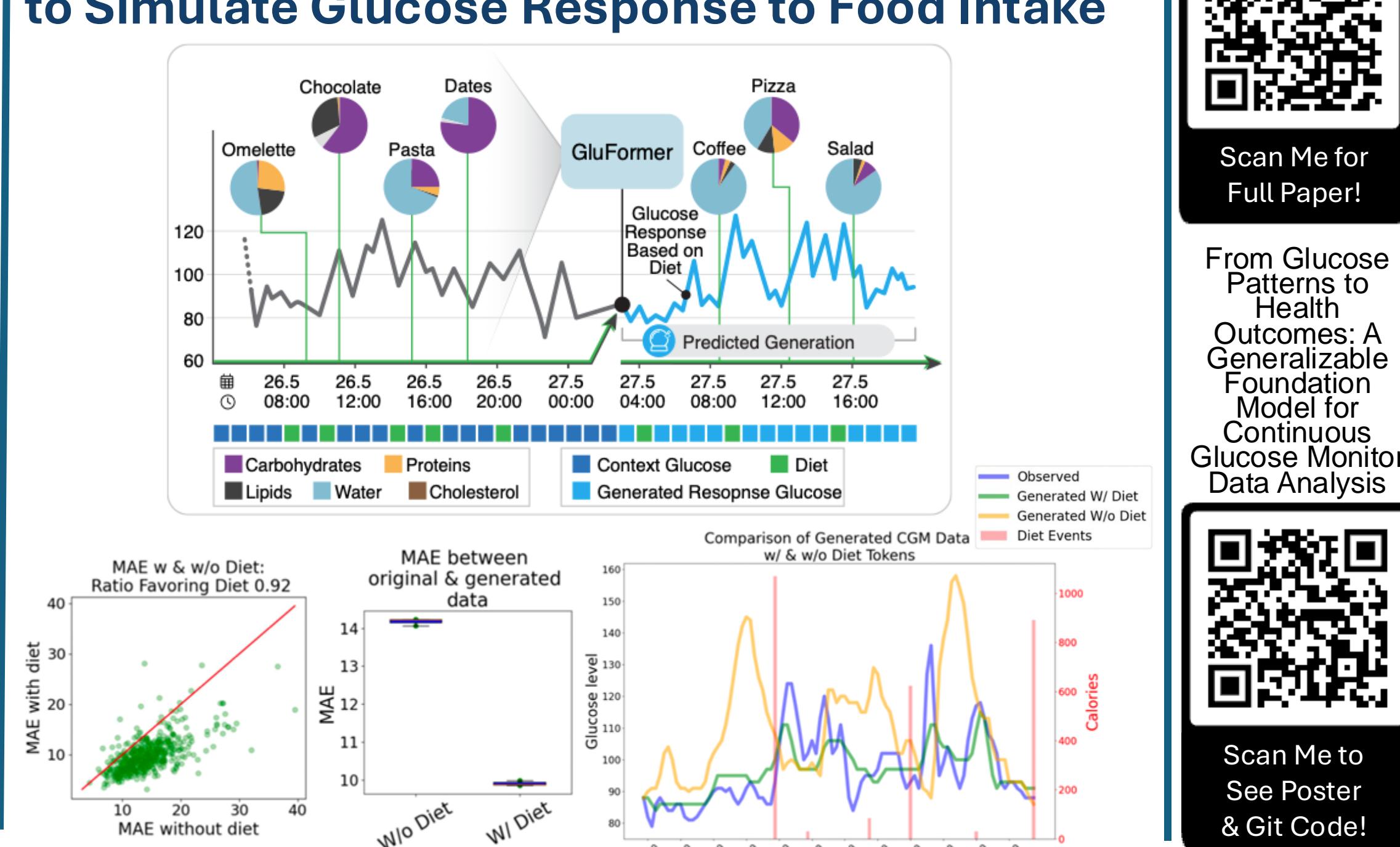
## 7. Predicted Hba1c% Acts as a Continuous 1-d Score to Rank Individuals on a Disease Progression Spectrum



## 8. GluFormer-derived Score is More Predictive of Future Diabetes than Gold-Standard Blood Test for Diabetes



## 9. Multi-Modal GluFormer: Adding Diet Loggings to Simulate Glucose Response to Food Intake



Scan Me for Full Paper!

From Glucose Patterns to Health Outcomes: A Generalizable Foundation Model for Continuous Glucose Monitor Data Analysis



Scan Me to See Poster &amp; Git Code!