Project Summary Report

YODA Project Protocol #: 2024-0248

Time-series insights into diabetes treatment – using a fine-tuned CGM foundation model to improve treatment outcomes

Objective
We aimed to develop and evaluate machine-learning methods that use baseline continuous glucose monitoring (CGM) time-series data to predict individual treatment response to anti-diabetic agents — specifically the SGLT-2 inhibitor canagliflozin and the DPP-4 inhibitor sitagliptin — and to identify CGM-derived signatures distinguishing responders from non-responders. The intent was to combine descriptive CGM analyses (CGMap, IGLU) with a fine-tuned time-series foundation model trained on an external CGM cohort, then apply the resulting representations to participant-level data from the requested trials.

Methods
We accessed participant-level CGM data and supporting documentation for trials NCT03267576 (canagliflozin vs. sitagliptin in metformin-treated T2DM, crossover) and NCT02139943 (canagliflozin add-on to insulin in T1DM). Consistent with the original proposal, the analytical workflow began with data preparation, exploratory analysis, and computation of standard CGM summary metrics — mean glucose, time-in-range, glycemic variability, and glucose management indicator (GMI) — using the CGMap and IGLU pipelines. We then performed exploratory unsupervised projections (e.g., PCA, UMAP) of the resulting feature space, stratified by treatment arm and timepoint, and conducted between-arm comparisons of pre- and post-treatment CGM summaries. We did not complete the methodological extensions described in the original proposal — fine-tuning of the TimeGPT foundation model on external CGM data, training of supervised treatment-response classifiers, and per-cluster outcome analyses — because the preliminary findings described below indicated that the available data would not support meaningful downstream modeling.

Results
In the preliminary analyses, between-arm differences in standard CGM-derived metrics were small in magnitude and did not reach statistical significance under reasonable correction for multiple comparisons. Unsupervised projections of the CGM feature space did not reveal separable structure aligned with treatment arm or with on-treatment glycemic outcomes. Effect sizes for the CGM signals we examined were substantially below those needed to support the supervised treatment-response prediction tasks defined in the original proposal, and neither cohort offered sufficient sample size to compensate for this in a foundation-model fine-tuning regime. Given these findings, we did not proceed to TimeGPT fine-tuning or to model evaluation against held-out trial outcomes.

Conclusions
The accessed trials contained the variables specified in our original proposal, and the preliminary analyses were executed as planned. However, the magnitude of treatment-related CGM signal in these cohorts was insufficient to support the intended development of a fine-tuned CGM time-series foundation model for individual treatment-response prediction. We therefore did not complete the central modeling aims of the proposal, and no publishable findings were generated.