Using Early Engagement Data from a Digital Health Solution to Predict Future Health Outcomes

Junjie Luo, MS[2], Abhimanyu Kumbara, MS[1], Anand Iyer, PhD[1], Mansur Shomali, MD, CM[1], Gordon Gao, PhD[2]
[1]Welldoc, Columbia, Maryland,USA,[2]Center for Digital Health and Artificial Intelligence, Johns Hopkins Carey Business School, Baltimore, MD,USA

BACKGROUND/PURPOSE

Engagement with a digital health solution is required for it to deliver its intended purpose to an individual with a chronic condition such as diabetes. In diabetes, it is well established that optimal glucose management requires: 1. an understanding of an individual's glucose levels in real-time; and 2. diligent self-management of key behaviors including medication administration, diabetes education, diet, activity and lab data (MEDAL). The combination of continuous glucose monitors (CGM), which can provide dense glucose data, and AI-driven, regulated digital health solutions, which capture the MEDAL user engagement data, helps to address these two factors. More research is needed to understand how the combination of dense CGM data and MEDAL data allow for optimal glucose management.

SPECIFIC AIMS

We have previously shown that combining dense CGM and MEDAL engagement data can be used to accurately predict future MEDAL engagement.[1] In this new analysis, we hypothesized that early engagement signals from CGM wear time as well as MEDAL engagement data could help predict future health outcomes, specifically Time in Range (TIR) at days 70 to 90 after initial engagement.

METHODS

We reviewed real-world data from 499 individuals with type 1 (55%) and type2 (45%) diabetes, who were enrolled in a combined CGM-digital health solution program. We used the first 30 days from enrollment with the digital health solution as the baseline period, and days 70 to 90 from the enrollment date for the prediction period. In the prediction period, all individuals had to have at least 70% CGM data sufficiency (n=304). TIR >= 0.7 was labeled as a good health outcome, and TIR <0.7 was labelled as not a good health outcome. CGM features, MEDAL engagement features, and demographics at baseline period were used to predict the binary TIR outcome variable in the prediction period. Prediction models were run for 4 sub-group populations: Overall, Type 2 Diabetes only (T2D), Type 1 Diabetes only (T1D), and CGM features only.

Figure 1: Screenshots of Digital Health Solution

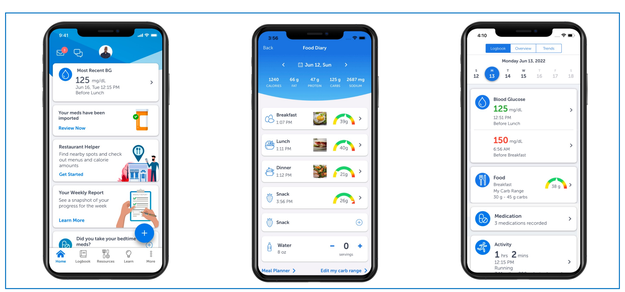
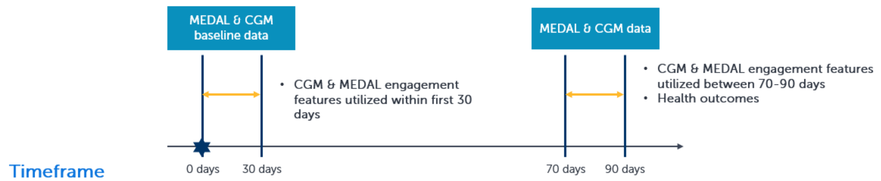


Figure 2: Baseline Period and Outcome Period Timeframe



RESULTS

Figure 3: Demographic Information

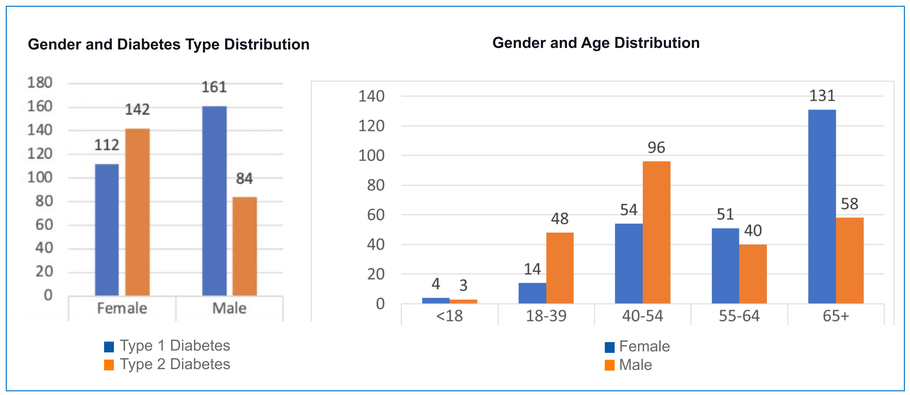


Figure 4: Population CGM Metrics

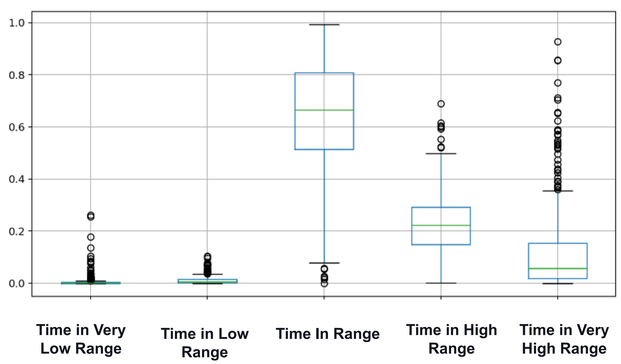


Figure 5: Prediction Model Output Statistics for Each Sub-group Analysis Population

Question Variants	Data (n)	Best Model	Accuracy	AUC	Precise	Recall	F1
Q1 Baseline	304	Light Gradient Boosting Machine	0.7970	0.8813	0.8248	0.7500	0.7812
Q2: T2D	140	Random Forest Classifier	0.7956	0.8740	0.8198	0.7700	0.7833
Q3: T1D	125	Extra Trees Classifier	0.8403	0.9050	0.8539	0.8850	0.8504
Q4: No MEDAL	304	Light Gradient Boosting Machine	0.8206	0.8799	0.8528	0.7700	0.8035

- We observed that the female group in this research sample had a higher proportion of T2D when compared to that of males, who had a higher proportion of T1D.
- The female group in the research sample was also relatively skewed towards an older age when compared to that of the male population.
- A LightGBM (gradient boost model) predicted the health outcome variable in the overall population with ~80% accuracy, with an AUC score of 0.88.
- Other sub-groups also exhibited similar accuracies, with the T1D sub-group having the highest model prediction accuracy at ~84%, with an AUC score of 0.9.
- Across all subgroups, some form of tree-based prediction model was shown to be best suitable for predicting future TIR.

CONCLUSIONS

- Early CGM data, along with MEDAL engagement data, can be useful in predicting health outcomes such as TIR (Time In Range) in future periods of an individual's journey.
- Prediction models, such as the ones described in this study, can be integrated into future digital health solutions to improve the personalization and precision that digital health coaching can provide to a person living with a chronic condition.
- Factors such as gender and age can be instrumental in building predictive models for specific micro-segments of the population.
- Future research will explore, with the help of different CGM outcome variables, the specific user engagement behaviors that contribute most to optimal glucose management.

REFERENCES

[1]Junjie Luo, Abhimanyu Kumbara, Anand Iyer, Mansur Shomali, Gordon Gao. Using Early Engagement Data from a Digital Health Solution to Predict Future Engagement Patterns. Presented at Conference on Health IT and Analytics (CHITA), Washington DC, May 5-6, 2023.