TITLE: Using LLMs for automated narrative generation from diabetes device data to support low-burden data review by clinicians

PROBLEM: Type 1 diabetes (T1D) is a complex condition associated with severe health complications, hospitalizations, and premature mortality[1]. Its lifelong management requires continuous review of complex data from diabetes devices, such as Continuous Glucose Monitors (CGM) and insulin pumps, that are available via commercial device data platforms[2–5]. Current digital platforms impose a significant cognitive and temporal burden on clinicians, limiting effective review of data in short clinic visits, resulting in clinician burnout and reduced patient engagement[6,7]. Ineffective engagement with data despite available data platforms results in missed opportunities for care and suboptimal blood glucose control[8,9]. Using Large Language Models (LLMs) to augment diabetes device data review has the potential to transform how data is utilized as it can reduce manual data analysis by automating insight generation[10], thereby lowering the burden of data review. Additionally, LLM-based modules can integrate with telemedicine platforms to support low-burden data review at scale, positively impacting patient care and accessibility for underserved populations worldwide. Lastly, mechanisms to maintain transparency by providing clinicians with the necessary context to trust and validate AI-generated insights can ensure they are ethically and responsibly integrated into patient care. In this project, we aim to use LLMs to generate comprehensive data narratives from multidimensional diabetes device data and explore the use of these narratives by clinicians in reviewing data to make care decisions.

PRELIMINARY DATA: Our prior work has shown that different diabetes device data streams (blood glucose, bolus insulin, basal insulin, meal events, and automated pump events) when presented as multimodal narratives lower the burden of data review and improve data comprehension and decision-making among T1D patients[11]. These narratives are detailed explanations of glycemic episodes, contextualized with potential causal factors and supporting data, to provide a comprehensive understanding of the patient's condition. Given the promise of narratives in supporting low-burden engagement with multidimensional data, we will develop automated narrative generation module using GPT-4o and real patient data along with an interface to present these narratives. For this, we have continuous glucose monitor and insulin pump data from several patients from past studies[12]. Additionally, we have the PDF reports from these data that will be used by clinicians to review data and generate narratives that serve as ground truth.

SPECIFIC AIMS: Our aims are: (1) Identify clinicians' information needs and questions in working with diabetes device data (2) Develop and test prompts for GPT to generate data-based narratives and visuals tailored to clinicians' questions and information needs (3) Compare GPT and clinician performance on narrative generation task.

MILESTONES: (1) November 2024–February 2025: For specific aim 1, we will conduct focus groups and interviews with endocrinologists to determine information sought by clinicians, such as notable glycemic episodes, insulin use, and patient behaviors. (2) March 2025–June 2025: For specific aim 2, we will select patient cases from our past study data, create prompts for querying GPT-4o, compare GPT-4o's performance in generating narratives with raw device data only and with raw device data along with statistical features, and conduct initial tests using ground truth provided by clinical experts. (3) July 2025–October 2025: For specific aim 3, we will have 5 clinicians generate narratives, compare AI-generated and clinician-generated narratives based on parameters such as accuracy, usefulness, and time efficiency, and gather qualitative feedback from 20 clinicians in one-hour interviews to understand the usability and accuracy of AI-generated narratives.

BUDGET: Graduate student research assistant ($32320), human subjects study compensation for 20 clinicians ($100/hour, $2000), student software developer ($6000), OpenAI credits (~$200/month, $2400 for 12 months), and conference attendance for PI and student ($6500). Total budget: $49,220.

TEAM: Shriti Raj, PhD (Assistant Professor of Medicine, Center for Biomedical Informatics Research); David Maahs, MD, PhD (Lucile Salter Packard Professor of Pediatrics, Endocrinology and Diabetes); Vasilis Syrgkanis, PhD (Assistant Professor of Management Science and Engineering); Hee Jung Choi (Graduate Student Research Assistant, Computer Science)

COLLABORATION: Our prior work, access to real patient data, and an interdisciplinary team with expertise in machine learning, human-computer interaction, and endocrinology positions us to successfully undertake this project. Involvement of machine learning and human-computer interaction experts will help us not only understand the feasibility of generating computational outputs but also deliver those outputs in ways that align with user needs. Involvement of endocrinology experts will help us design ethical, holistic, and augmentative AI-based solution that integrates with existing clinical practices to help improve health outcomes.

REFERENCES:

1. The Effect of Intensive Treatment of Diabetes on the Development and Progression of Long-Term Complications in Insulin-Dependent Diabetes Mellitus. *N Engl J Med*. 1993;329(14):977-986. doi:10.1056/NEJM199309303291401

2. Ng SM, Sultana P, Clemente M, Apperley L. Comparison Between Patients and Families who Routinely Download Data and Those who do Not Download Data at Home in The Management of Type 1 Diabetes. In: *ESPE Abstracts*. Vol 92. Bioscientifica; 2019. Accessed January 19, 2024. https://abstracts.eurospe.org/hrp/0092/hrp0092p1-314

3. Lee JM, Rusnak A, Garrity A, et al. Feasibility of Electronic Health Record Assessment of 6 Pediatric Type 1 Diabetes Self-management Habits and Their Association With Glycemic Outcomes. *JAMA Netw Open*. 2021;4(10):e2131278. doi:10.1001/jamanetworkopen.2021.31278

4. Hinnen DA, Buskirk A, Lyden M, et al. Use of Diabetes Data Management Software Reports by Health Care Providers, Patients With Diabetes, and Caregivers Improves Accuracy and Efficiency of Data Analysis and Interpretation Compared With Traditional Logbook Data. *J Diabetes Sci Technol*. 2014;9(2):293-301. doi:10.1177/1932296814557188

5. Irace C, Schweitzer MA, Tripolino C, Scavelli FB, Gnasso A. Diabetes Data Management System to Improve Glycemic Control in People With Type 1 Diabetes: Prospective Cohort Study. *JMIR MHealth UHealth*. 2017;5(11):e170. doi:10.2196/mhealth.8532

6. Tanenbaum ML, Adams RN, Lanning MS, et al. Using Cluster Analysis to Understand Clinician Readiness to Promote Continuous Glucose Monitoring Adoption. *J Diabetes Sci Technol*. 2018;12(6):1108-1115. doi:10.1177/1932296818786486

7. Wong JC, Neinstein AB, Spindler M, Adi S. A Minority of Patients with Type 1 Diabetes Routinely Downloads and Retrospectively Reviews Device Data. *Diabetes Technol Ther*. 2015;17(8):555-562. doi:10.1089/dia.2014.0413

8. Ebekozien O, Mungmode A, Sanchez J, et al. Longitudinal Trends in Glycemic Outcomes and Technology Use for Over 48,000 People with Type 1 Diabetes (2016-2022) from the T1D Exchange Quality Improvement Collaborative. *Diabetes Technol Ther*. 2023;25(11):765-773. doi:10.1089/dia.2023.0320

9. Foster NC, Beck RW, Miller KM, et al. State of Type 1 Diabetes Management and Outcomes from the T1D Exchange in 2016–2018. *Diabetes Technol Ther*. 2019;21(2):66-72. doi:10.1089/dia.2018.0384

10. Healey E, Tan A, Flint K, Ruiz J, Kohane I. Leveraging Large Language Models to Analyze Continuous Glucose Monitoring Data: A Case Study. *medRxiv*. Published online April 8, 2024:2024.04.06.24305022. doi:10.1101/2024.04.06.24305022

11. Raj S, Gupta T, Lee J, Kay M, Newman MW. “It can bring you in the right direction”: Episode-Driven Data Narratives to Help Patients Navigate Multidimensional Diabetes Data to Make Care Decisions. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI ’23. Association for Computing Machinery; 2023:1-16. doi:10.1145/3544548.3581073

12. Prioleau T, Bartolome A, Comi R, Stanger C. DiaTrend: A dataset from advanced diabetes technology to enable development of novel analytic solutions. *Sci Data*. 2023;10(1):556. doi:10.1038/s41597-023-02469-5