

Modelo Fundacional *Zero-Shot* e Dados Multimodais para Predição Probabilística de Glicemia em Diabetes Tipo 1

Mario Cesar Freire Dias Filho¹, Rodrigo de Azeredo Siqueira²
 Vinicius Silva Osterne Ribeiro¹, Danielo G. Gomes¹

¹Programa de Pós-Graduação em Engenharia de Teleinformática (PPGETI)
 Centro de Tecnologia, Universidade Federal do Ceará (UFC) – Fortaleza – CE

²Universidade Iguaçu (UNIG) – Nova Iguaçu – RJ

mariocesarfreire@alu.ufc.br, vinicius@osterne.com
 rodrigoendocrinologista@gmail.com, danielo@ufc.br

Abstract. *Blood glucose prediction is a key challenge in type 1 diabetes (T1D) management, as glycemic levels result from dynamic interactions between insulin, nutrition, physical activity, and physiological factors. Here we show that the Chronos-2 foundation model overcomes the need for patient-specific training by providing zero-shot glycemic forecasts with a high proportion of predictions in clinically acceptable error zones. We evaluated the model's performance on the OhioT1DM dataset for horizons of 30 to 120 min using multimodal covariates. The model kept 98.45% of predictions in clinically acceptable zones (A+B) of the Clarke Error Grid at 30 min, with 92.33% within the hypoglycemic sub-region. While covariates failed to reduce point accuracy errors, they enhanced uncertainty calibration and brought coverage closer to the nominal level. Consequently, foundation models like Chronos-2 establish a robust probabilistic baseline for automated T1D management, pending clinical validation.*

Resumo. *A predição da glicose no sangue é um desafio central no manejo do diabetes tipo 1 (DM1), pois os níveis de glicemia resultam de interações dinâmicas entre insulina, alimentação, atividade física e fatores fisiológicos. Neste artigo, mostramos que o modelo fundacional Chronos-2 supera a necessidade de treinamentos específicos ao oferecer previsões glicêmicas com elevada fração em zonas de erro clinicamente aceitáveis em regime zero-shot. Avaliamos o desempenho do modelo no conjunto OhioT1DM para horizontes de 30 a 120 minutos, utilizando covariáveis multimodais. O modelo manteve 98,45% das predições nas zonas clinicamente aceitáveis (A+B) da Clarke Error Grid. As covariáveis não reduziram o erro pontual de forma consistente, mas melhoraram a calibração da incerteza (P10-P90), aproximando a cobertura do nível nominal. Assim, modelos fundacionais como o Chronos-2 estabelecem um novo patamar de baseline probabilístico robusto, ainda sujeito a validação clínica, para o manejo automatizado do DM1.*

1. Introdução

O diabetes tipo 1 (DM1) exige monitoramento contínuo da glicose (*Continuous Glucose Monitoring*, CGM) e decisões terapêuticas frequentes para manter a glicemia em faixas

seguras. Oscilações rápidas podem causar hipoglicemia ou hiperglicemia, ambas associadas a riscos clínicos relevantes [Limbert et al. 2024]. A predição glicêmica tornou-se, assim, um componente importante do suporte à decisão, antecipando correções de insulina, ingestão de carboidratos e ajustes de atividade física. Como a dinâmica glicêmica resulta da interação entre múltiplos fatores fisiológicos e comportamentais (da aplicação de insulina à qualidade do sono), trata-se de um problema inerentemente multimodal e temporal.

Abordagens supervisionadas têm explorado essa multimodalidade com resultados promissores, incluindo arquiteturas convolucionais temporais [Santos et al. 2025], *Transformers* personalizados por paciente [Sun et al. 2026] e estratégias híbridas com modelagem metabólica e nutricional [Wolber et al. 2025], mas normalmente dependem de treinamento específico no domínio, personalização individual ou dados supervisionados em larga escala.

Recentemente, modelos fundacionais de séries temporais passaram a ser investigados como alternativas mais gerais para cenários com disponibilidade limitada de dados [Moor et al. 2023], com resultados iniciais sugerindo desempenho competitivo em predição glicêmica de curto horizonte [Rancati et al. 2024]. Entre os modelos fundacionais recentes, o Chronos-2 [Ansari et al. 2025] destaca-se por operar em regime *zero-shot*, isto é, realizando inferência com pesos congelados, usando o histórico de cada paciente apenas como contexto de entrada, sem qualquer treinamento, *fine-tuning* ou ajuste de gradiente no domínio de destino.

Neste artigo, avaliamos o Chronos-2 em regime *zero-shot* para predição glicêmica probabilística em DM1 utilizando covariáveis exclusivamente históricas (*past-only*). Investigamos como diferentes modalidades fisiológicas influenciam simultaneamente o erro pontual e a calibração probabilística das previsões. Além disso, propomos um protocolo temporal replicável e reproduzível no OhioT1DM [Marling and Bunescu 2020] e posicionamos o desempenho do Chronos-2 frente a *baselines* supervisionados reportados para a mesma família de dados.

2. Material e Método

2.1. Conjunto de Dados

Usamos dados do OhioT1DM [Marling and Bunescu 2020], com 17 participantes (10 mulheres e 7 homens, entre 23 e 70 anos) acompanhados por aproximadamente 12 semanas em condições reais de vida. O conjunto reúne leituras de CGM com frequência aproximada de 5 min, registros de insulina basal e *bolus*, ingestão de carboidratos, proteínas, gorduras e fibras, métricas de atividade física e dados de sono [Broussard et al. 2016].

2.2. Pré-processamento e Normalização sem Vazamento

As séries foram alinhadas em uma grade temporal de 5 min: variáveis contínuas foram interpoladas ou acumuladas, e eventos discretos (insulina, refeições) foram representados como impulsos temporais. Aplicou-se escalonamento robusto (mediana e IQR) [Rousseeuw and Hubert 2011] com parâmetros estimados exclusivamente nas 10 semanas iniciais de cada participante e aplicados sem modificação às 2 semanas finais de avaliação, evitando vazamento temporal. As saídas foram desnormalizadas para avaliação em mg/dL.

2.3. O Modelo Chronos-2 e o Regime Zero-Shot

O Chronos-2 [Ansari et al. 2025] é um modelo fundacional pré-treinado para previsão universal de séries temporais, com suporte nativo a séries univariadas e multivariadas e a covariáveis *past-only* e *known-future*. A inferência é probabilística e baseada em aprendizado em contexto, com pesos congelados e sem ajuste de gradiente em novos domínios. Neste artigo, *zero-shot* significa inferência com parâmetros congelados: o histórico de cada participante serve apenas como contexto, sem atualização de pesos, *fine-tuning* ou calibração supervisionada. Todas as covariáveis são tratadas em modo *past-only*, reproduzindo um cenário realista em que refeições, exercícios e sono futuros são desconhecidos. O processamento usou GPU NVIDIA A100 (40 GB), em lotes por indivíduo.

2.4. Cenários de Entrada e Protocolo de Avaliação

A estratégia experimental combinou horizontes de 30, 60 e 120 minutos, cada um com janela de contexto de mesma duração (isto é, o contexto fornecido ao modelo tem exatamente o tamanho do horizonte previsto), e quatro cenários incrementais de covariáveis: A1 (apenas glicose), A2 (glicose e insulina), A3 (glicose, insulina e macronutrientes) e A4 (glicose, insulina, nutrição, atividade física e sono) [Cappon et al. 2019, Turksoy et al. 2015]. Validamos de forma temporal *hold-out* [Bergmeir and Benítez 2012]: as primeiras 10 semanas de cada participante formaram o histórico acessível e as previsões foram geradas em janelas deslizantes nas 2 semanas finais. A escolha das 10 semanas é uma decisão de protocolo para reservar *hold-out* temporal e maximizar o contexto disponível; a sensibilidade a janelas menores permanece em aberto.

2.5. Métricas de Avaliação

Reportamos MAE e RMSE em mg/dL para precisão pontual, calculados ponto a ponto sobre o *pool* global de predições dos 17 participantes (Tabela 1); a Tabela 2 complementa com a dispersão interindividual de MAEs computados por participante. A calibração probabilística é avaliada pela cobertura do intervalo P10-P90, isto é, a fração de valores reais entre os quantis preditivos de 10% e 90%, idealmente próxima de 0,80. Os quantis são produzidos nativamente pelo Chronos-2 por inferência probabilística em contexto, sem amostragem Monte Carlo, e a cobertura é agregada ponto a ponto sobre o mesmo *pool* global. Para utilidade clínica, complementamos com a *Clarke Error Grid Analysis* (CEG), que mapeia pares (real, previsto) em cinco zonas: A e B agregam predições clinicamente aceitáveis; C, D e E representam erros potencialmente perigosos. Reportamos a fração em A+B e na sub-região hipoglicêmica (Hypo A+B), relevante por ser a hipoglicemia o evento de maior risco agudo em DM1.

3. Resultados

3.1. Desempenho Pontual e Probabilístico Agregado

A Tabela 1 resume MAE, RMSE e cobertura P10-P90. A avaliação agregada compreende $N = 46.088$, 46.032 e 45.909 previsões para os horizontes de 30, 60 e 120 min, respectivamente, somando os pontos de *hold-out* com alvo CGM real dos 17 participantes. Em 30 min, o MAE varia de 8,67 a 8,93 mg/dL entre A1 e A4 (diferenças marginais); em 60 e 120 min o crescimento é monotônico e suave. A cobertura P10-P90 cresce consistentemente com a multimodalidade, ultrapassando 0,80 no cenário A4 e aproximando-se do nível nominal.

Tabela 1. Métricas de desempenho pontual (MAE/RMSE) e calibração probabilística (P10-P90) por cenário de covariáveis e horizontes de predição.

Cenário	MAE (mg/dL)			RMSE (mg/dL)			Cob. P10-P90		
	30	60	120	30	60	120	30	60	120
A1 (Unimodal)	8,67	14,52	23,08	13,32	22,25	34,57	0,780	0,776	0,775
A2	8,86	14,74	23,29	13,45	22,45	34,83	0,781	0,775	0,777
A3	8,89	14,78	23,38	13,48	22,52	34,96	0,793	0,786	0,784
A4 (Multimodal)	8,93	14,81	23,40	13,55	22,57	34,92	0,813	0,806	0,803

Tabela 2. Desempenho do cenário A4 por horizonte: zonas da Clarke Error Grid e dispersão interindividual do MAE para os 17 participantes.

Horizonte	Clarke Error Grid (%)			Dispersão do MAE (mg/dL)				
	A+B	Hypo A+B	C+D+E	Mediana	Q1	Q3	Min	Max
30 min	98,45	92,33	1,55	8,71	7,42	10,28	5,96	12,54
60 min	95,12	86,40	4,88	14,38	12,11	17,05	9,80	21,13
120 min	89,67	78,10	10,33	22,94	19,40	27,21	15,82	34,90

3.2. Avaliação Clínica e Dispersão Interindividual (Cenário A4)

A Tabela 2 reúne, para o cenário A4 multimodal e por horizonte, a distribuição clínica pela *Clarke Error Grid* e a dispersão interindividual do MAE. Em 30 min, 92,33% das predições caem em A+B na sub-região hipoglicêmica (Hypo A+B), a mais sensível para segurança em DM1, e a degradação em horizontes longos acompanha o crescimento do erro pontual. O IQR do MAE também cresce com o horizonte, indicando maior heterogeneidade entre participantes em previsões de longo prazo.

4. Discussão

A Tabela 3 apresenta a comparação interna entre o ARIMA e o Chronos-2 no cenário A1 (glicose univariada), ambos avaliados sob protocolo idêntico: *split* fixo de 10 semanas de treino e 2 semanas de *hold-out* por participante, escalonamento robusto (mediana/IQR) ajustado exclusivamente no período de treino e avaliação em mg/dL sobre o mesmo *pool* de 17 participantes do OhioT1DM. O ARIMA foi ajustado via AIC apenas nos dados de treino, sem *refit* no período de teste.

O Chronos-2 *zero-shot* apresenta MAE e RMSE consistentemente menores que o ARIMA em todos os horizontes, com vantagem de MAE crescente ($\approx 9,8$ mg/dL a 30 min até $\approx 17,6$ mg/dL a 120 min). Sob protocolo idêntico para ambos, isso evidencia que o pré-treinamento massivo do Chronos-2 captura padrões de autocorrelação glicêmica além de um modelo autorregressivo linear, sobretudo em horizontes longos. A vantagem estende-se à calibração: sua cobertura P10–P90 (0,780 a 30 min) aproxima-se mais do nominal de 0,80 que a do ARIMA (0,720), cujos intervalos gaussianos paramétricos subcobrem a incerteza, distinção explorada na Seção 4.

Para contextualização adicional, a Tabela 4 posiciona o Chronos-2 frente a resultados publicados na literatura. Como coortes, datasets e protocolos diferem, essas cifras

Tabela 3. Desempenho no protocolo OhioT1DM: ARIMA univariado comparado ao Chronos-2 A1 *zero-shot*. Destaques apontam o melhor resultado por horizonte.

Modelo	Estratégia	Horizonte	MAE (mg/dL)	RMSE (mg/dL)	Cob. P10–P90
ARIMA	univariado · AIC · OhioT1DM	30 min	18,42	23,84	0,720 [‡]
		60 min	27,33	38,57	0,735 [‡]
		120 min	40,68	53,29	0,719 [‡]
Chronos-2 (A1)	<i>zero-shot</i> · OhioT1DM	30 min	8,67	13,32	0,780
		60 min	14,52	22,25	0,776
		120 min	23,08	34,57	0,775

[‡]Intervalos gaussianos paramétricos (IC 80%), nominalmente equivalentes ao P10–P90, mas de natureza distinta dos quantis nativos do Chronos-2. ■ Melhor resultado por horizonte.

Tabela 4. Posicionamento frente à literatura (trabalhos com diferentes coortes e datasets). Destaques indicam o melhor desempenho por janela.

Artigo	Modelo · Estratégia · Dataset	Janela de pred.	MAE (mg/dL)	RMSE (mg/dL)
[Santos et al. 2025]	Vanilla TCN · sup. · n/disp. [†]	15 min	4,91	8,44
	Vanilla TCN · sup. · n/disp. [†]	90 min	13,84	22,94
[Rancati et al. 2024]	TimeGPT · zero-shot · n/disp. [†]	30 min	15,59	18,26
[Wolber et al. 2025]	mLLM+mec. · sup. · AZT1D	30 min	–	16,61
	mLLM+mec. · sup. · AZT1D	60 min	–	24,58
	mLLM+mec. · sup. · AZT1D	120 min	–	31,22
[Sun et al. 2026]	Teacher · fut.-aware · OhioT1DM	30 min	12,73	17,83
	Teacher · fut.-aware · OhioT1DM	60 min	21,50	28,81
	Teacher · fut.-aware · OhioT1DM	120 min	32,42	41,93
	Chronos-2 · zero-shot · OhioT1DM	30 min	8,93	13,55
Este artigo	Chronos-2 · zero-shot · OhioT1DM	60 min	14,81	22,57
	Chronos-2 · zero-shot · OhioT1DM	120 min	23,40	34,92

[†]Dataset proprietário. ■ Melhor resultado por janela de predição.

valem como referência de ordem de grandeza e não como comparação controlada (papel da Tabela 3). Ainda assim, o Chronos-2 *zero-shot* situa-se próximo ao Vanilla TCN supervisionado [Santos et al. 2025] em horizontes médios e tem erro menor em 30 min que o TimeGPT [Rancati et al. 2024]; notavelmente, *student* e *teacher* de [Sun et al. 2026], sobre a mesma família OhioT1DM, apresentam MAE/RMSE maiores que o Chronos-2 em todos os horizontes.

4.1. Multimodalidade: Calibração mais que Erro Pontual

A multimodalidade incremental de A1 a A4 não reduz MAE/RMSE de forma consistente, mas melhora a cobertura P10-P90, particularmente em 30 min (Tabela 1). Em regime *zero-shot*, sem mapeamentos aprendidos entre carboidratos e elevação glicêmica nesta coorte, as covariáveis aumentam a sensibilidade do modelo à incerteza do contexto, refinando a dispersão prevista sem deslocar a estimativa central. Isso não implica irrelevância: a representação como impulsos preserva o instante e a composição da refeição, mas não modela atrasos de absorção. Modelagem explícita desses atrasos via curvas de

absorção por macronutriente [Wolber et al. 2025] ou compartimentos farmacocinéticos de IOB e COB [Sun et al. 2026] pode ampliar o efeito sobre a estimativa central.

A comparação com o ARIMA esclarece a natureza dos ganhos de cobertura. Com bandas gaussianas paramétricas, o ARIMA atinge apenas 0,720 a 30 min (abaixo do nominal de 0,80), evidenciando subcobertura. O Chronos-2 em A1, com quantis nativos, já cobre mais próximo do nominal (0,780), e a multimodalidade (A4) eleva a cobertura a 0,813. Como o Chronos-2 parte de intervalos mais estreitos e ainda assim cobre melhor que a banda gaussiana do ARIMA, o ganho de cobertura reflete refinamento da incerteza condicional, não mero alargamento dos intervalos.

4.2. Aplicabilidade Clínica

A interpretação dos horizontes deve considerar a utilidade de decisão: 30 min apoiam alerta de risco iminente; 60 min sustentam ajustes de ingestão; 120 min servem mais à antecipação de risco que à ação pontual de alta precisão [Limbert et al. 2024]. A fração de 92,33% em Hypo A+B em 30 min (Tabela 2) é especialmente relevante por cobrir a janela mais sensível a hipoglicemia. A combinação de erro moderado, cobertura próxima ao nominal e alta concentração em zonas aceitáveis sugere que o modelo, mesmo *zero-shot*, pode operar como *baseline* probabilístico clinicamente informado.

Ressalta-se que os resultados se baseiam em análise retrospectiva por zonas de erro (*Clarke Error Grid*) e não constituem validação clínica: o modelo não foi avaliado em cenário prospectivo de apoio à decisão nem em administração automática de insulina. As afirmações sobre adequação clínica restringem-se à concordância com zonas de erro aceitáveis nas condições experimentais descritas.

4.3. Viabilidade Computacional e de Implantação

Os dois modelos diferem em recursos. O ARIMA, leve, ajusta-se por participante em dezenas de segundos em CPU, com inferência de 10–12 s por duas semanas de *hold-out* e latência desprezível, sendo compatível com execução local em dispositivos de baixo custo. O Chronos-2 (120 milhões de parâmetros) é mais exigente (≈ 18 min em CPU para os 17 participantes), mas seu *footprint* de memória ($\approx 480/240$ MB em precisão simples/reduzida) cabe em *smartphones* modernos, embora a latência em dispositivo permaneça em aberto. Como o CGM gera uma leitura a cada 5 min, a implantação favorece dois modos: inferência local com o ARIMA sob hardware restrito e inferência remota com o Chronos-2 quando há conectividade, transmitindo o contexto recente a um servidor, o que dispensa GPU no dispositivo. A inviabilidade ocorreria apenas sem conectividade e sem hardware local para o modelo fundacional, quando o ARIMA permanece como alternativa, ainda que com erro maior.

5. Conclusão

A predição glicêmica em DM1 permanece desafiadora pela natureza multimodal e dinâmica da glicose em vida livre. Mostramos que o Chronos-2 em regime *zero-shot* produz previsões estáveis e clinicamente consistentes sem treinamento no domínio ou personalização por paciente, alcançando 98,45% das previsões em A+B da *Clarke Error Grid* em 30 min e calibração intervalar próxima ao nominal sob incorporação multimodal.

Os resultados também indicam um desacoplamento parcial entre erro pontual e calibração: a multimodalidade trouxe ganhos limitados nas métricas pontuais, mas melhorou a confiabilidade probabilística, contribuindo mais para a quantificação de incerteza do que para a redução do erro médio. Sob protocolo idêntico, o Chronos-2 superou o *baseline* ARIMA em erro pontual e calibração em todos os horizontes, sugerindo que modelos fundacionais de séries temporais capturam padrões glicêmicos relevantes com forte generalização, mesmo em coortes reduzidas e sem adaptação ao domínio.

Essas limitações apontam direções de trabalho futuro conectadas aos *gaps* atuais. A coorte reduzida de 17 participantes motiva estratégias leves de personalização, capazes de capturar a resposta glicêmica individual sem abrir mão da generalidade *zero-shot*. A representação das covariáveis como impulsos simples, que limitou o ganho de multimodalidade no erro pontual, motiva a modelagem farmacocinética explícita (IOB/COB), capaz de deslocar a estimativa central e não só refinar a incerteza. A limitação de amostra motiva transformação e aumento de dados [Alexandrino et al. 2024] como complemento ao pré-treinamento. Por fim, avaliar um único modelo em um único regime motiva a comparação sistemática entre modelos fundacionais, isolando o efeito de arquitetura do de pré-treinamento.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Danielo G. Gomes agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de produtividade (processos 309014/2026-3, 311845/2022-3).

Disponibilidade de Dados e Código

O conjunto OhioT1DM [Marling and Bunesu 2020] é disponibilizado pelos autores originais mediante solicitação formal por e-mail (vide <https://webpages.charlotte.edu/rbunesu/data/ohiot1dm/OhioT1DM-dataset.html>). Os scripts de pré-processamento e inferência utilizados neste artigo podem ser solicitados aos autores via e-mail.

Referências

- Alexandrino, F., Pacheco, C., Carvalho, D., and Ogasawara, E. (2024). Aumento de dados e suavização integrada para predição de séries temporais baseada em aprendizado de máquina. In *Anais do XVIII Brazilian e-Science Workshop (BreSci 2024)*, pages 32–39. Sociedade Brasileira de Computação.
- Ansari, A. F., Shchur, O., Küken, J., Auer, A., Han, B., Mercado, P., Rangapuram, S. S., Shen, H., Stella, L., Zhang, X., Goswami, M., Kapoor, S., Maddix, D. C., Gueron, P., Hu, T., Yin, J., Erickson, N., Desai, P. M., Wang, H., Rangwala, H., Karypis, G., Wang, Y., and Bohlke-Schneider, M. (2025). Chronos-2: From univariate to universal forecasting.
- Bergmeir, C. and Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213.
- Broussard, J. L., Ehrmann, D. A., Van Cauter, E., Tasali, E., and Brady, M. J. (2016). Sleep restriction impairs insulin sensitivity in human adipocytes. *Diabetologia*, 59:359–368.

- Cappon, G., Vettoretti, M., Sparacino, G., and Facchinetti, A. (2019). Continuous glucose monitoring sensors for diabetes management: a review of technologies and applications. *Diabetes & Metabolism Journal*, 43(4):383–397.
- Limbert, C., Kowalski, A. J., and Danne, T. P. (2024). Automated insulin delivery: A milestone on the road to insulin independence in type 1 diabetes. *Diabetes Care*, 47(6):918–920.
- Marling, C. and Bunesco, R. (2020). The OhioT1DM dataset for blood glucose level prediction: Update 2020. In *Proceedings of the 5th International Workshop on Knowledge Discovery in Healthcare Data (KDH@ECAI 2020)*, volume 2675 of *CEUR Workshop Proceedings*, pages 71–74. CEUR-WS.org.
- Moor, M., Banerjee, O., Abad, Z. S. H., Krumholz, H. M., Leskovec, J., Topol, E. J., and Rajpurkar, P. (2023). Foundation models for generalist medical artificial intelligence. *Nature*, 616(7956):259–265.
- Rancati, S., Bosoni, P., Schiaffini, R., Deodati, A., Mongini, P. A., Sacchi, L., Toffanin, C., and Bellazzi, R. (2024). Exploration of foundational models for blood glucose forecasting in type-1 diabetes pediatric patients. *Diabetology*, 5(6):584–599.
- Rousseeuw, P. J. and Hubert, M. (2011). Robust statistics for outlier detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):73–79.
- Santos, M., Áurea Melo, Santos, J., Gonçalves, I., Santos, W., and R., L. (2025). Modelagem de séries temporais biomédicas: Um estudo comparativo entre psta-tcn e tcn vanilla para previsão de dados de glicose de pessoas diabéticas. In *Anais do XXV Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 991–996, Porto Alegre, RS, Brasil. SBC.
- Sun, X., Li, H., and Yu, X. (2026). Future-aware blood glucose forecasting using knowledge distillation with transformer-based sequence-to-sequence models. *Scientific Reports*, 16:11404.
- Turksoy, K., Bayrak, E. S., Quinn, L., Littlejohn, E., and Cinar, A. (2015). Multivariable adaptive identification and control for artificial pancreas systems. *IEEE Transactions on Biomedical Engineering*, 63(3):526–534.
- Wolber, J., E. Samadi, M., Sellin, J., and Schuppert, A. (2025). Multimodal large language models and mechanistic modeling for glucose forecasting in type 1 diabetes patients. *Journal of Biomedical Informatics*, 172:104945.