

Comparing Deterministic and Stochastic Reinforcement Learning for Glucose Regulation in Type 1 Diabetes

David Timms¹, Dr Chirath Hettiarachchi¹, Prof Hanna Suominen^{1,2}

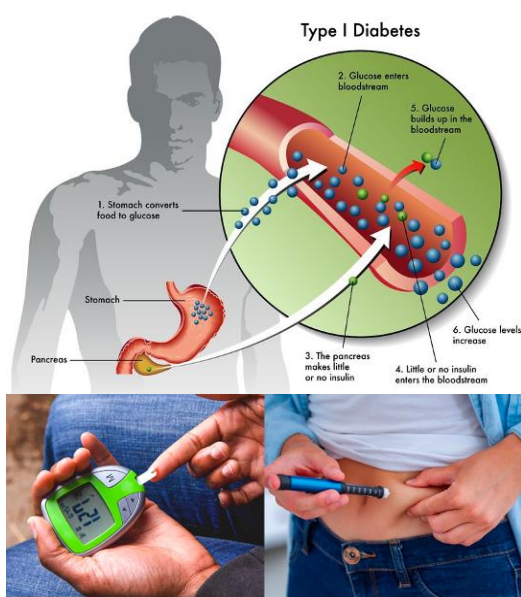
¹The Australian National University, Australia

²University of Turku, Finland

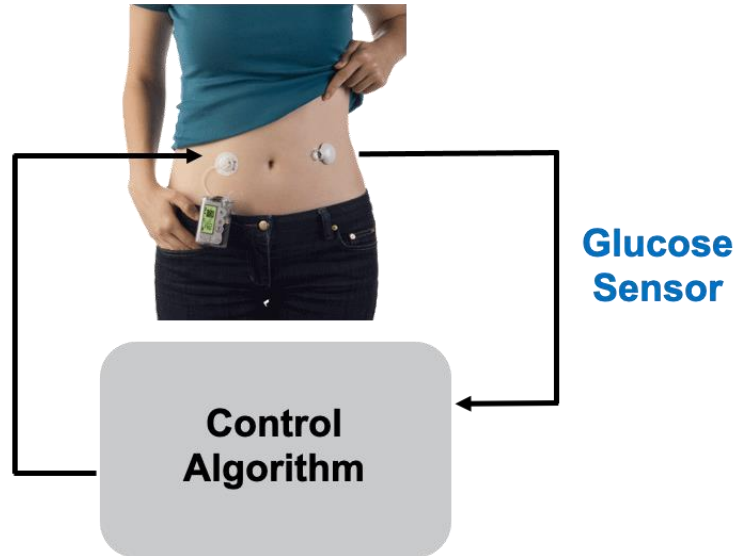


Australian
National
University

Background: Glucose Regulation in Type 1 Diabetes (T1D)



External insulin Administration is required!



Artificial Pancreas Systems (APS)
 Existing APS are hybrid-closed loop systems



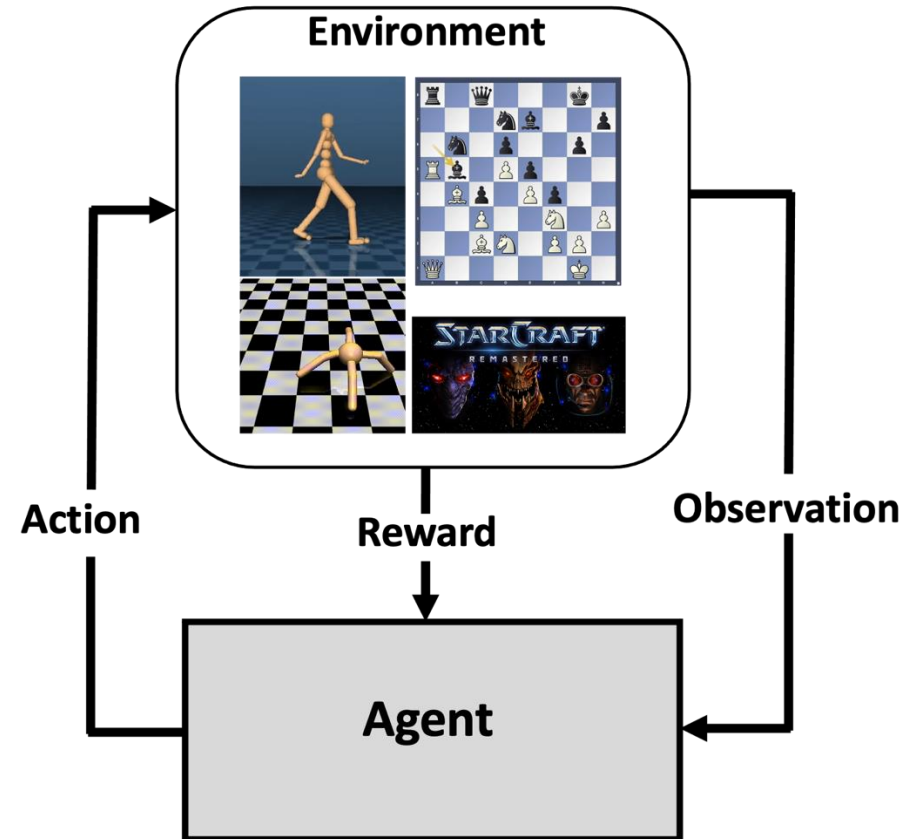
Manual meal announcements & carbohydrate estimation is required.

[1] DiMeglio, Linda A., Carmella Evans-Molina, and Richard A. Oram. "Type 1 diabetes." *The Lancet* (2018).

[2] Diabetes Control and Complications Trial Research Group. "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus." *New England Journal of Medicine* (1993).

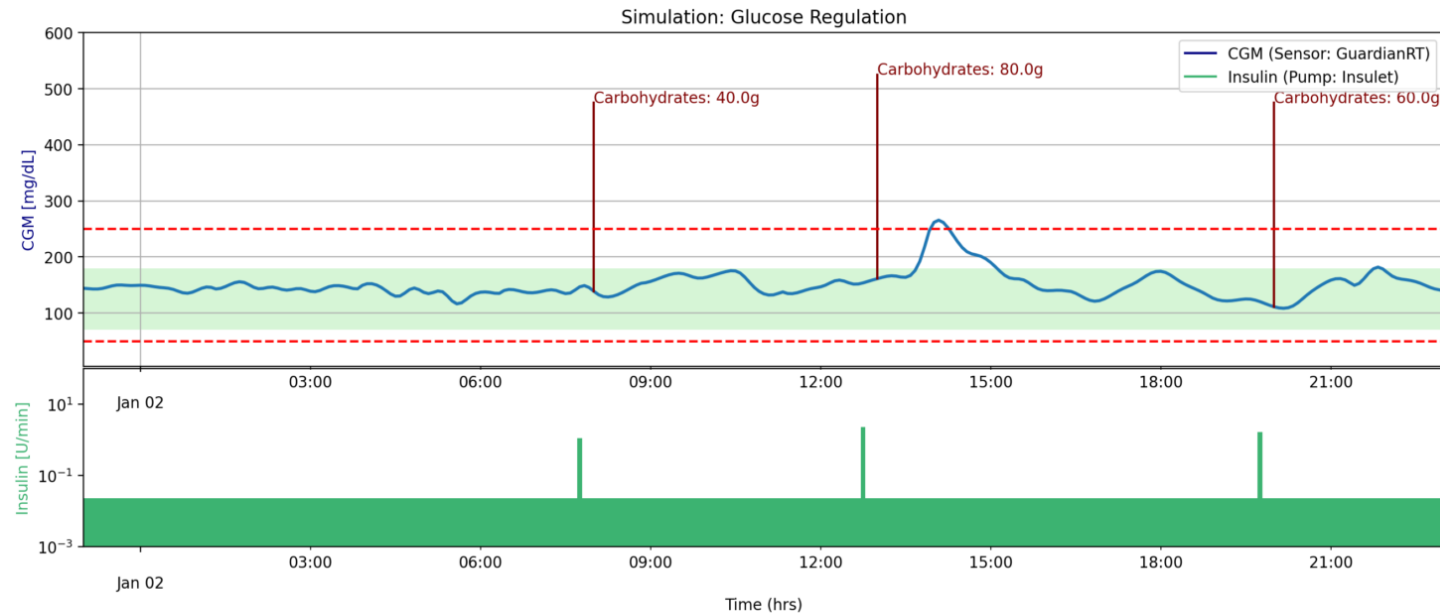
[3] Hettiarachchi, et al. "A reinforcement learning based system for blood glucose control without carbohydrate estimation in type 1 diabetes: In silico validation." *44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (2022).

Background: Reinforcement Learning (RL)



Introduction

In T1D treatment, the clinical objective is to improve Time in Normoglycemic Range (TIR), while avoiding hypoglycemic & hyperglycemic risk [1, 2].

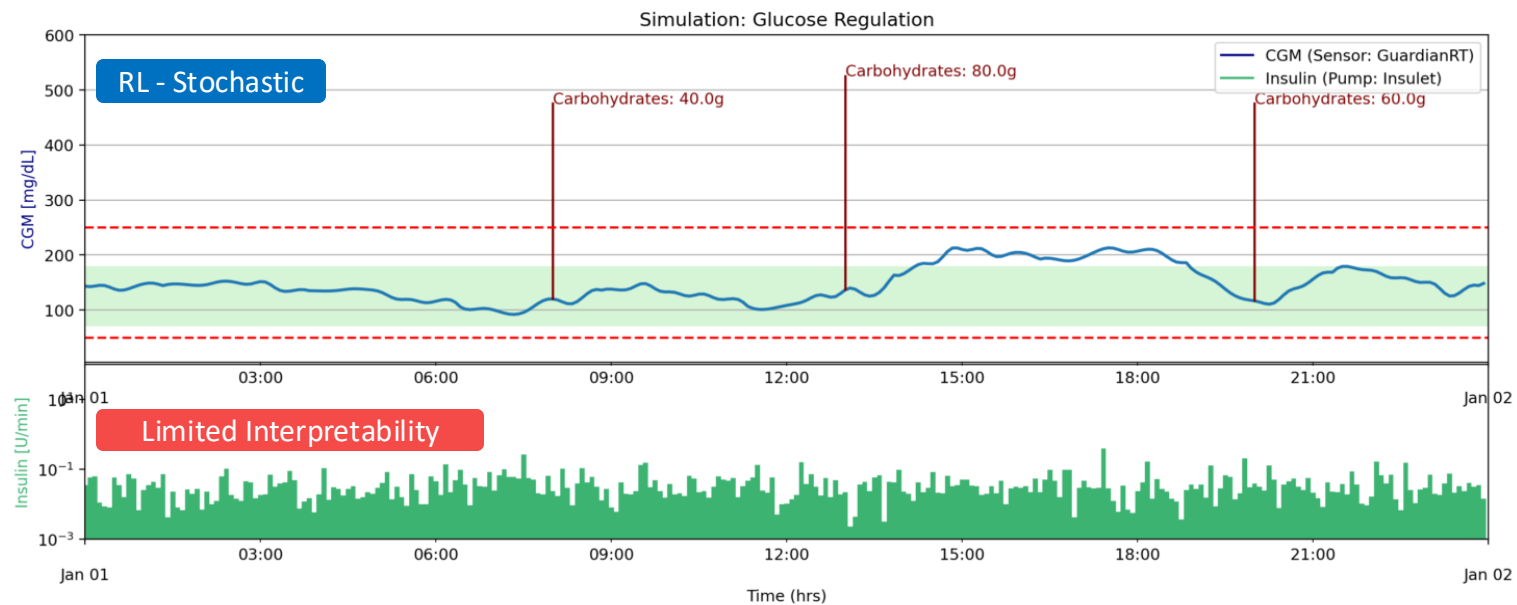


Clinical treatment strategy (BBI [2])

[4] Hettiarachchi, et al. "G2P2C—A modular reinforcement learning algorithm for glucose control by glucose prediction and planning in Type 1 Diabetes." *Biomedical Signal Processing and Control* (2024).
 [5] Emerson, et al. "Offline reinforcement learning for safer blood glucose control in people with type 1 diabetes." *Journal of Biomedical Informatics* (2023).
 [6] Lee, et al. "Toward a fully automated artificial pancreas system using a bioinspired reinforcement learning design: In silico validation." *IEEE Journal of Biomedical & Health Informatics* (2020).

Study Aim

Recent research is exploring RL for developing treatment strategies aimed at improving glucose regulation performance and automating manual interventions (e.g., meal announcements) [3-6].

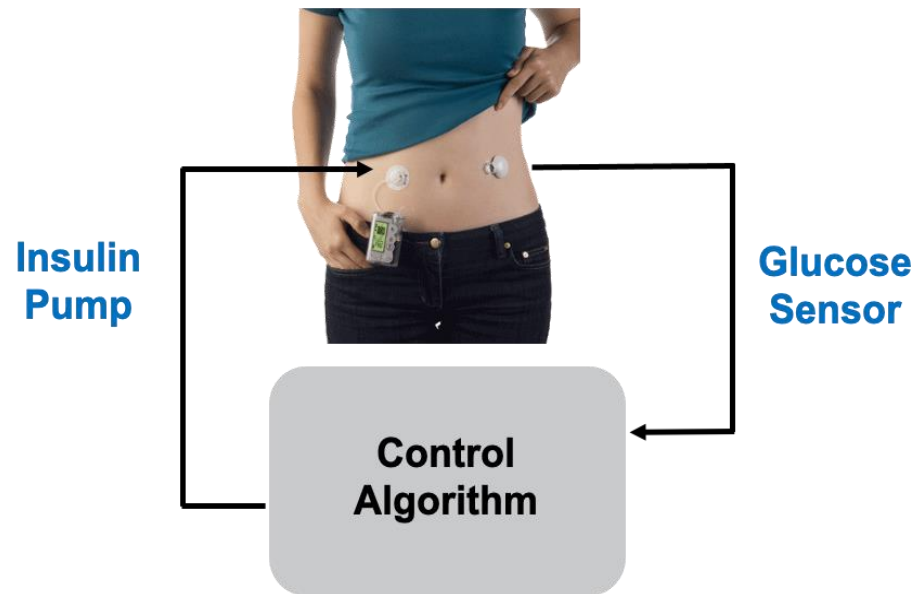


RL-based strategy (G2P2C [4])

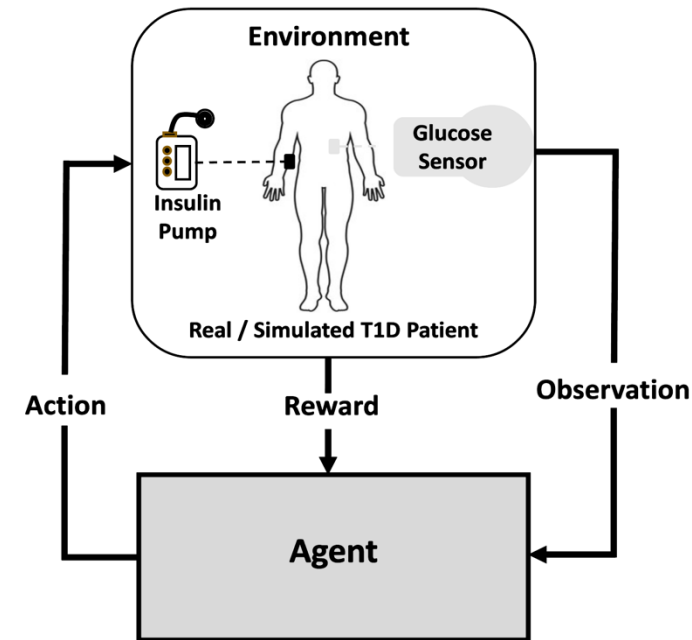
Previous work is predominantly based on "stochastic" RL algorithms; which limits "interpretability" (critical for safety). Therefore, we explore "deterministic" RL and conduct a comparative analysis.

Methods: Experiment Setup

- APS are **high-risk medical devices**. Following established best practices for pre-clinical trials, we use the FDA-accepted **UVA/PADOVA 2008** model [7] in the open-source Simglucose simulator for reproducibility [8].
- We simulate dynamics of real-world T1D populations, by selecting **in-silico patients** (10 adults/adolescents), **commercial pumps/sensors**, and **meal schedules** (breakfast, lunch, dinner; variable times/carbs with uncertainty in occurrence) [3].



Commercial Artificial Pancreas Systems (APS)



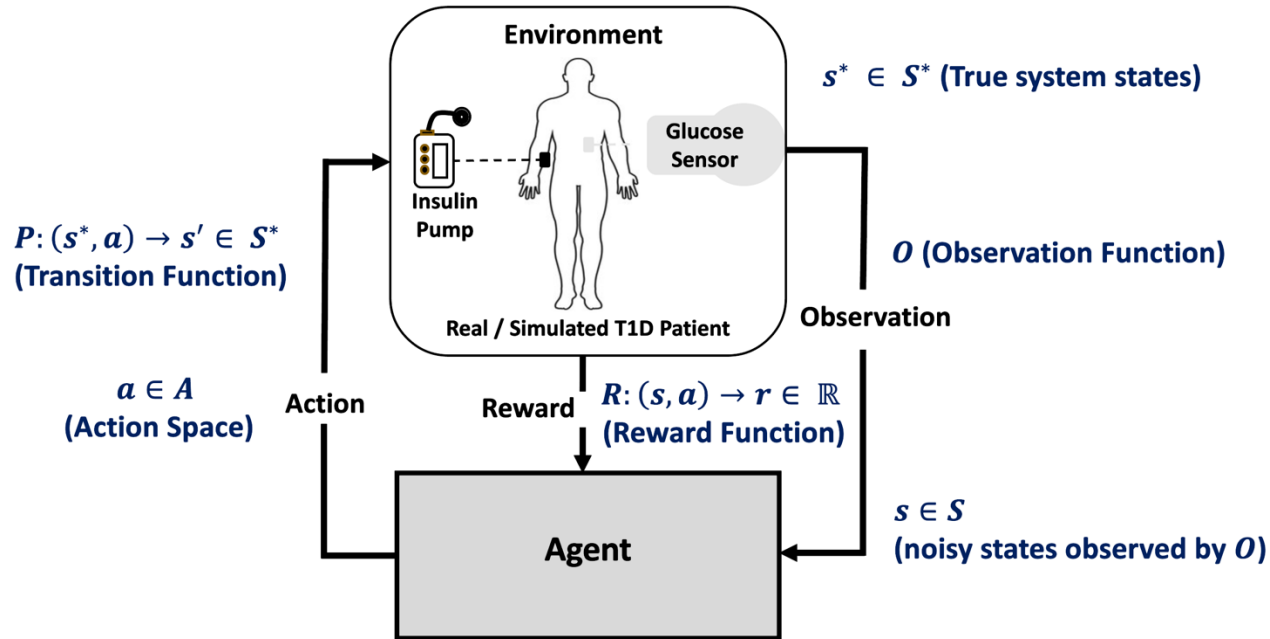
Experimental Simulations/Setup

[7] Kovatchev, Boris P., et al. "In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes." *Journal of Diabetes Science and Technology* (2009).

[8] Xie, Jinyu, 2018. Simglucose v0.2.1. <https://github.com/jxx123/simglucose>.

Methods: Problem Formulation & RL/Clinical Algorithms

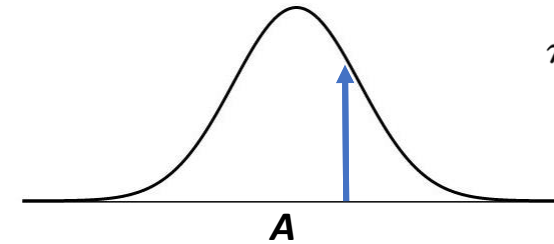
(A) RL task formulated as a POMDP. Detailed information: [3, 5, 9]



(B) Candidate Algorithms for glucose control.

(1) Proximal Policy Optimisation (PPO)

RL - Stochastic

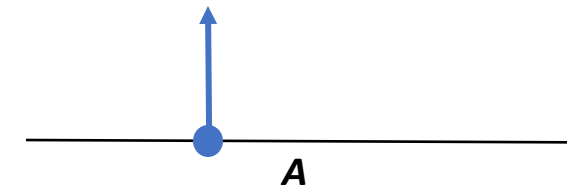


$$\pi(a | s) = \mathbb{P}[A = a | S = s]$$

$$a \sim \pi(\cdot | s)$$

(2) Twin Delayed Deep Deterministic Policy Gradient (TD3)

RL - Deterministic



$$a = \mu(s)$$

(3) Basal Bolus Ideal (BBI)

Clinical

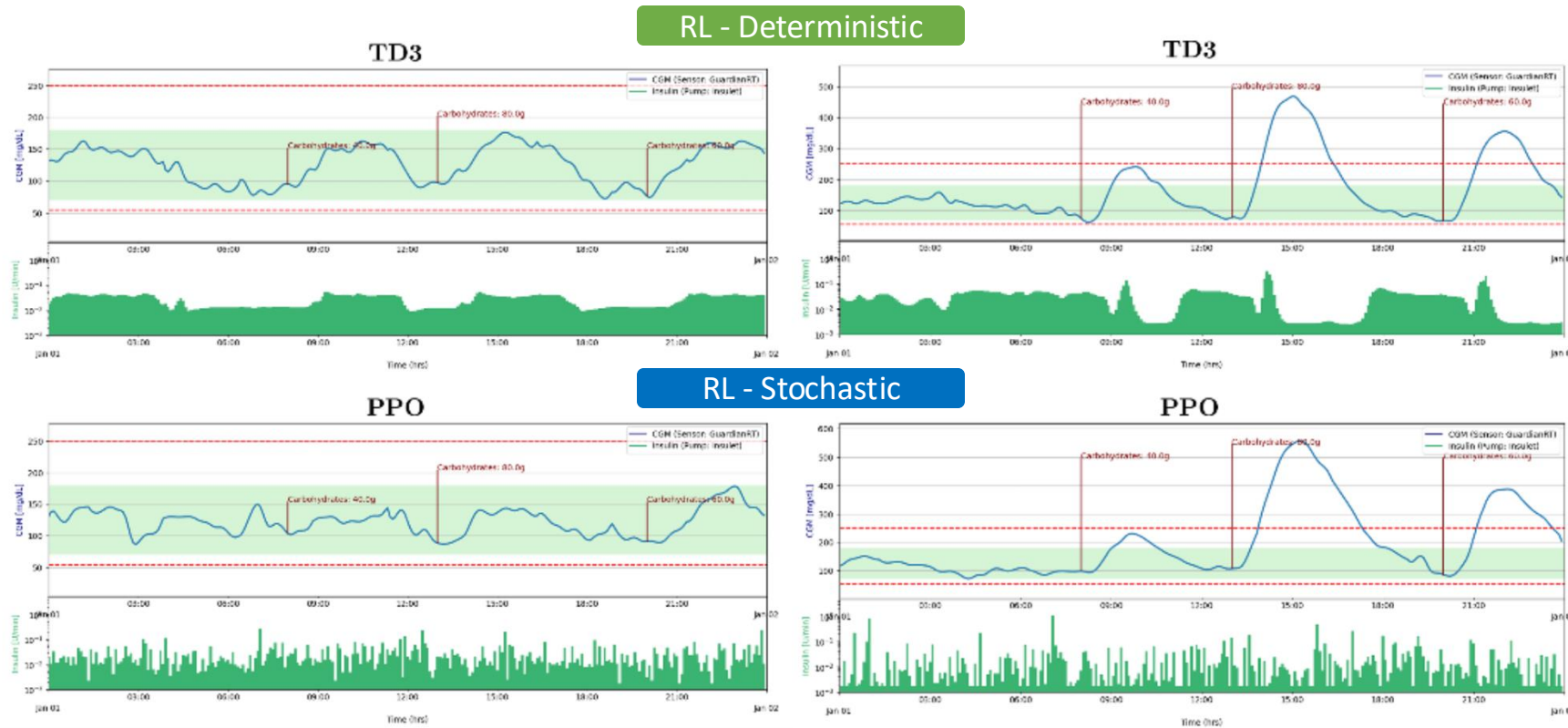
Meal estimation/announcement

A: rule-based, based on patient-specific parameters.

[9] Hettiarachchi, et al. "Non-linear continuous action spaces for reinforcement learning in type 1 diabetes." *Australasian Joint Conference on Artificial Intelligence* (2022).

Results

- **PPO (stochastic) is the best performing RL algorithm.** However, its insulin response is hard to interpret.
- In contrast, **TD3 (deterministic) shows smoother insulin administration and is more predictable.**
- For TD3, we identified that more recent blood glucose & insulin data have a higher correlation with administered insulin.



Results

- PPO and TD3 did not outperform the clinical benchmark BBI in Time In normoglycemic Range (TIR) & Failure Rate (FR).
- However, **BBI requires perfect meal announcements & carbohydrate estimation information**, which was not provided to the RL algorithms.

Table 1. Performance comparison of PPO, TD3, and BBI algorithms

Algorithm	Adolescents (TIR / FR)	Adults (TIR / FR)
BBI	71.43 ± 12.31 / 0%	71.02 ± 11.29 / 0.39%
PPO	63.72 ± 13.95 / 4.93%	69.12 ± 10.53 / 2.79%
TD3	60.99 ± 19.18 / 25.6%	62.79 ± 16.30 / 18.37%

* The evaluation meal protocol spanned 24 hours starting at 00:00hrs fixed with three meals: 40g of CHO for breakfast at 8:00hrs, 80g of CHO for lunch at 13:00hrs, and 60g of CHO for dinner at 20:00hrs.

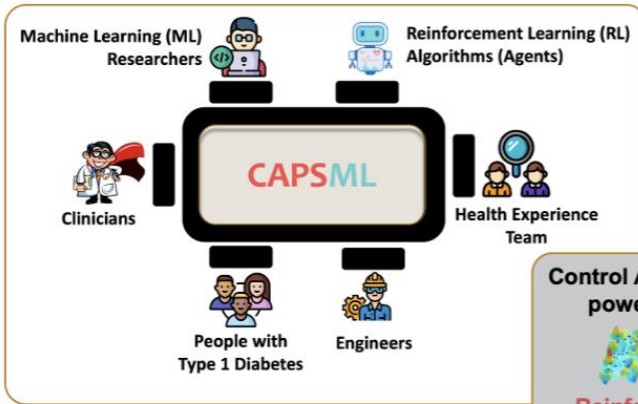
Conclusion

- Although the behaviour of TD3 is easier to interpret; and achieves a smoother insulin administration; **it limits the adaptability of the algorithm to quickly adjust insulin delivery in response to sudden changes in blood glucose levels**, which can challenge effective regulation.
- This conclusion challenges assessing algorithmic safety and suitability, also **highlighting the importance of improving APS applications for both interpretability and predictive performance in future research.**
- **Limitations:** This study only focused on *in-silico* adult/adolescent populations and future research should focus on extending the analysis to the child cohort.

Open-source Resources

CAPSML

1,900+ users,
56 countries.
(July 2023 – Aug 2025)



Alice
Age: 18, Gender: F, Weight: 68.7Kg
Total Daily Insulin (TDI): 36.73
Insulin to Carbohydrate Ratio (ICR): 12
Insulin Sensitivity Factor (ISF): 15.03

Select *in-silico* subjects

Control Algorithms powered by
Reinforcement Learning

Learn about your meals
How much carbs are in a McDonalds Cheese Burger?
How does it affect glucose?



Case Study I: Clinical Treatment vs ML Methods
Chloe is an *in-silico* adult with T1D. We want to compare how her blood glucose is controlled by a basal-bolus clinical treatment and an AI strategy named G2P2C.

Case Study II: Meal Variability
David is an *in-silico* adult with T1D. David was late to work and forgot to have his breakfast. Let's find out how it would affect his glucose when an AI strategy is used for glucose control.

Case Study III: Patient Variability
Alice & Bob are *in-silico* adolescents with T1D. We want to compare their variability in glucose control using an AI strategy named G2P2C while the meal protocol is kept fixed.



RL4H
Reinforcement Learning for Health
Repositories: G2P2C, GluCoEnv, RL4T1D

☆ 54
👤 27

GluCoEnv

GluCoEnv - Glucose Control Environment is a simulation environment which aims to facilitate Reinforcement Learning based Artificial Pancreas Systems for Glucose Control in Type 1 Diabetes.

About
This project implements *in-silico* Type 1 Diabetes (T1D) subjects for developing glucose control systems. The project aims to facilitate the development of Reinforcement Learning (RL) based Artificial Pancreas Systems (APS) for the glucose control problem. The figure below shows the main components of an APS.

Installation & Dependencies
Create a Python 3.10.12 virtual environment and install PyTorch 2.2.0.

```
python3 -m venv env
source env/bin/activate
pip install --upgrade pip
```

The project can be installed using the source or pip. To install using the source,

```
git clone https://github.com/chirathy/glucoenv.git
cd glucoenv
pip install -e .
```

G2P2C: Reinforcement Learning based Artificial Pancreas Systems.

Background: Type 1 Diabetes (T1D) is caused by the autoimmune destruction of the islet beta-cells and results in absolute insulin deficiency (cover image: Human islet of Langerhans created by Stable Diffusion). Hence, external administration of insulin is required to maintain glucose levels, which is crucial as both low and high glucose levels are detrimental to health. This is usually done through an insulin pump attached to the body. An artificial pancreas system (APS) is a closed-loop system that automates the insulin administration process. In this project we design Reinforcement Learning (RL) based Artificial Pancreas Systems (APS) for the glucose control problem. The figure below shows the main components of an APS.

Maintaining glucose levels is a life-long optimization problem, complicated due to the disturbances associated with daily events (meals, exercise, stress, etc.), delays present in glucose sensing and insulin action, observability, and safety constraints among others. A simulation of glucose regulation, using an RL-based APS, is shown below, where the optimal glucose range is shaded in green (severe hypoglycemia / hyperglycemia highlighted by the red dotted line). The blood glucose measurements are presented in the top, while the administered insulin by the RL agent is presented in the bottom. The disturbances related to meal and carbohydrate content of the meals are presented in red.

RL4T1D: Reinforcement Learning for Automating Treatment in Type 1 Diabetes.

Background: Type 1 Diabetes (T1D) is caused by the autoimmune destruction of the islet beta-cells and results in absolute insulin deficiency (cover image: Human islet of Langerhans created by Stable Diffusion). Hence, external administration of insulin is required to maintain glucose levels, which is crucial as both low and high glucose levels are detrimental to health. This is usually done through an insulin pump attached to the body. An artificial pancreas system (APS) is a closed-loop system that automates the insulin administration process. In this project we design Reinforcement Learning (RL) based Artificial Pancreas Systems (APS) for the glucose control problem. The figure below shows the main components of an APS.

Visit: capsml.com

Visit: <https://github.com/RL4H>

Acknowledgement

- This research was delivered in partnership with Our Health in Our Hands, a strategic initiative of ANU, which aims to transform health care by developing new personalized health technologies and solutions in collaboration with patients, clinicians, and health-care providers.
- We gratefully acknowledge funding from the MRFF 2022 National Critical Research Infrastructure (MRF-CRI000138, Developing a new digital therapeutic for depression: Closed loop non-invasive brain stimulation).
- This work was supported by computational resources provided by the Australian Government through the National Computational Infrastructure under the ANU Merit Allocation Scheme (ny83 and eu59) and ANU Startup Scheme (sj53).



Australian
National
University

