

# Comparison of PID controller, basal bolus controller, and reinforcement based solutions for insulin suggestions for type 1 diabetics

## I. INTRODUCTION

In 2019 over 463 million people worldwide were estimated to suffer from diabetes mellitus and numbers are rising year by year estimated to reach 700 million in 2045 [1] [2]. The disease not only affects individuals and their families but is a burden on society with an estimated global health expenditure of 727 billion USD [1]. The recommendations for good diabetes management state that the average person with type 1 diabetes (pwT1D) should be in the target range for more than 70% of the day, less than 25% in hyperglycemia (with less than 5% in severe hyperglycemia), and less than 4% in hypoglycemia (with less than 1% in severe hypoglycemia) [3]. The definition of the ranges might differ and are explained in detail in chapter II-C. Treatments for diabetes may vary depending on diabetes type, where all pwT1D and some with type 2 diabetes are insulin dependent. Insulin is one of the few medications where the administered dosage is decided and adapted by the patients themselves. Miscalculation or estimation of the dose to be injected might lead to severe hypo- or hyperglycemic events which can lead to death or permanently harm the patient. This task is further complicated by the complex system of blood glucose (BG) homeostasis in combination with the pharmacokinetics and glucodynamics of insulin. Therefore, one goal of diabetes technology is to adapt the dosage automatically or provide a suggestion on the amount of insulin to deliver.

Currently, no system is completely closed-looped. Commercially available systems currently only include a hybrid loop in which the basal insulin is automatically regulated but the meal boluses still have to be initiated by the user him- or herself [4]. Current approved systems include insulin suspension in case the continuous glucose monitoring (CGM) measures a value below a certain threshold, different target BG during exercise, and predictive low-glucose insulin suspension of basal insulin [4] [5] [6]. Algorithms used in research or commercial hybrid loops include Proportional-Integral-Derivative (PID) controller [5], model predictive controller (MPC) [7], fuzzy logic controller [4], or a combination of MPC and PID for bihormonal

systems. In this system both insulin and glucagon or another hormone with BG elevating function are used [8] [9].

This project aims to compare the performance of different controllers for automated control of BG levels in simulated T1D patients. The control of the BG only by controlling basal insulin is challenging. It is hypothesized that the use of reinforcement learning (RL) will yield a higher % of time in range (TIR) compared to the % TIR when using a PID controller. However, the basal bolus (BB) controller is expected to still outperform the PID and proximal policy optimization (PPO) controller.

In this project, different configurations of a PPO controller and a PID controller were compared to the performance of a BB controller. To this end, the established T1D simulator simglucose for the simulation as well as an implementation of a PPO controller in python were used.

## II. MATERIALS AND METHODS

### A. Setup

For this project, the BG of patients was simulated using simglucose [10] a T1D simulator for RL. The simulator is implemented in python and was used with the OpenAI gym [13] API. The simulation subjects included 30 patients with 10 adolescents, 10 adults, and 10 children. The scenario was chosen to have four meals per day with the following time and meal size in grams of carbohydrates (CHO): 7:00 (70g), 10:00 (30g), 14:00 (110g), 21:00 (90g). For each classifier the scenario was tested twice, once with and once without variability. The variability was introduced by randomly sampling a normal distribution centered around 0. For the meal times the standard deviation was set as  $\pm 15\text{min}$ . The meal size was varied by  $\pm 10\text{g}$  for the main meals and  $\pm 5\text{g}$  for the snack (10:00). Insulin sensitivity, correction factors and CHO to insulin ratios were created by the scenario by considering each subjects' own profile. The closed loop BG regulation can be formulated as an infinite-state Markov decision process with the parameters state, physiological model, action, reward function  $R(s_t, a_t)$  and discount factor. Summarized as  $\langle S, P, A, R, \gamma \rangle$ . The physiological model describes the transition from one

state to the next and is part of the simglucose simulator. The other parameters depend on the controller chosen for the BG control. Please refer to table I for the possible values of each controller. The different controller and their policy is described in the following subsection II-B. Since the simulator only accepts rates as insulin, both the basal and bolus insulin need to be calculated as rate. The insulin concentration is usually described in units (U) which corresponds to  $\frac{pmol}{L}$  in SI units. When insulin is mentioned in this project the unit is considered to be in  $\frac{U}{min}$  except if explicitly stated differently.

	BB	PID	PPO
State	CGM, CHO	CGM	CGM
Action	basal, bolus	basal	basal
Reward	NA	NA	partial negativity risk difference
			0.999
gamma	NA	NA	0.9
			0.5

Table I: Parameters of the infinite Markov decision process for the different controllers

### B. Controller

**BB controller:** In the BB controller the basal insulin is constant throughout the day and aims to stabilize the BG in a steady state (see equation 1).

$$basal = \frac{u2ss * BW}{6000} \quad (1)$$

The computation considers the body weight (BW) of the patient and the individual steady-state insulin rate (u2ss) which is divided by 6000 ( $1\frac{U}{min} = 6000\frac{pmol}{L}$ ). Every time, the patient eats something the bolus insulin is calculated. The applied dose considers the CHO to insulin ratio of the patient and only considers the correction factor (CF) if the BG is above the target as shown in equation 2.

$$bolus = \left( \frac{CHO_{meal} * t}{CHO_{ratio}} + \frac{BG_{current} - BG_{target}}{CF} \right) / t \quad (2)$$

As the BB controller does not consider the future states both reward and discount factor are not included.

**PID controller:** The three components of a PID controller can be investigated separately. Sometimes terms can be omitted, when they do not contribute anything to the control which will simplify the controller to a PI, PD, or P controller. In this project, the PID was tuned in an exploratory fashion and tested for the best configuration found. The proportional component of the controller considers the current BG, the integral part the time above and below range, and the derivative part considers the rate and direction of change of BG. The PID controller similar to the BB controller does not consider reward and discount factor.

**PPO controller:** A PPO controller is an algorithm which implements a policy gradient method. [11]. The PPO controller implemented in the Autonomous-Insulin-Infusion-Controller

[12] was used and modified for this project. Multiple parameters including environment, action space, policy, reward function, discount factor etc., can be adjusted independently. The action space can be chosen as discrete or continuous. If discrete the model selects actions from a predefined list whereas if continuous it selects actions between two boundary limits. Within this project, the lower and upper boundaries were set to 0 and 30. An important parameter to select is the policy. Two policies were tested, both based on long short-term memory (LSTM) policies with a multilayer perceptron (MLP) for feature extraction. LSTM policies are based on recurrent neural networks (RNN) but implement a solution to the vanishing gradient problem. The network implements an actor-critic model where the critic estimates the value function of the state and the actor updates the policy distribution according to this value. The value function is also called Q-function  $Q^\pi(s, a)$  and maximizes the expected reward  $r_t = R(s_t, a_t)$ . The optimal Q-function  $Q^*(s, a)$  (equation 3) is a recursive function since it includes Q-values from future states ( $s'$ ). Instead of iterating over all the successor states the PPO controller approximates  $Q^*(s, a)$  by the parameters ( $\theta$ ) of the network  $Q^*(s, a) \approx Q(s, a; \theta)$ .

$$Q^*(s, a) = E'_x [R(s, a) + \gamma \max_a' Q^*(s', a')] \quad (3)$$

Furthermore, different reward functions and discount factor can be selected. The reward functions used are the risk difference and partial negativity. In risk difference, the risk indices for the current state and the previous state are calculated, and their difference is used as a reward value. When choosing partial negativity as a reward function, values with a large deviation from the target value are penalized more harshly than values close to the target which means that these values are associated with a negative reward and values close to the target are associated with a positive reward. The discount factor determines how the RL agent handles rewards and dictates how far into the future the rewards should be considered. If  $\gamma = 0$ , the agent will be completely myopic and only learn about actions that produce an immediate reward. If  $\gamma = 1$ , the agent will evaluate each of its actions based on the sum of all of its future rewards. In this project, we tested  $\gamma$  with the values 0.999, 0.9, and 0.5.

### C. Evaluation

The results were observed qualitatively by looking at the BG developments for the different settings. Furthermore, statistical analysis was performed using R (version 2021, [14]). The BB controller was compared to both the PPO and PID controller as well as the PPO was compared to the PID controller. The comparisons were done based on the mean difference using a t-test with a significance level  $\alpha$  of 0.95. To test the hypothesis TIR needed to be implemented. However, judging performance on a single measure might not be appropriate. Therefore, a multitude of parameters have been implemented and will support testing the hypothesis. The values for time in range (TIR), time above range (TAR), time below range

(TBR), time in severe hyperglycemia (THI), and time in severe hypoglycemia (TLO) can be found in table II. The risk indices for low BG (LBGI) and high BG (HBGI) are calculated according to the symmetrized blood glucose measurement scale proposed by Kovatchev et al [15]. The implementation is done by scaling the individual BG readings before averaging over 1h (equation 4, with  $n$  being the number of readings per hour). Based on these averages the LBGI, and HBGI per hour (equation 5).

$$BG_h = \frac{1}{n} \sum_n (1.5 * \ln(BG)^{1.084} - 5.381) \quad (4)$$

$$LBGI = 10 * (BG_h < 0)^2 \quad (5)$$

$$HBGI = 10 * (BG_h > 0)^2$$

The overall risk index (RI) is calculated by adding the LBGI and the HBGI. To get the LBGI, HBGI, and RI for the whole time period the average is applied.

name	range
TIR	70 <= BG <= 180
TAR	180 < BG <= 250
TBR	50 <= BG < 70
THI	BG > 250
TLO	BG < 50

Table II: Ranges used for evaluation of BG control

### III. RESULTS

The following chapter discusses the most successful controllers found by the tuning process. Looking at each controller individually and comparing the performance with the varied scenario with the performance with the static scenario no large differences can be observed. Even though in some cases the difference was significant the absolute difference was small, as shown in table IV or graphically in figure 2. Therefore, this chapter only presents results for the scenario with variation. However, the results apply to both scenarios. The most important values are mentioned in the table III or the sections themselves directly, for complete results please refer to the tables IV and V in the appendix. Generally, it can be seen that the standard deviation (SD) is often very large for the overall results and gets smaller when only considering the individual groups, but even splitting the patients into groups substantial inter-patient variability is observed.

#### A. BB controller

The performance of the BB controller is considered the baseline. Looking at the plotted 24h interval for all patients (figure 3a) it can be seen that the BB controller fails to keep the glycemia in the desired range for a very long time. Especially if there are meal disturbances. During the night the BG falls below  $40 \frac{mg}{dl}$  on average although there are quite a few nights or patients that keep a stable BG throughout it, indicated by the grey lines above the standard deviation. Large excursions can be seen at around 9am, 4pm and 10pm. Dips into hypoglycemia are most frequent around 4am, 2pm, and from 7pm

		BB vs. PID	BB vs. PPO	PPO vs. PID
TIR	mean diff	42.9	25.6	17.3
	p-value	1E-11	2E-09	6E-07
TAR	mean diff	7.8	1.5	6.3
	p-value	1.0	0.75	1.00
TBR	mean diff	11.6	6.8	4.8
	p-value	1.0	1.0	1.00
THI	mean diff	-26.4	-6.2	-20.2
	p-value	0.001	0.01	0.002
TLO	mean diff	-36.0	-27.7	-8.3
	p-value	2E-05	2E-05	0.10
LBGI	mean diff	-40.7	-9.5	-31.1
	p-value	0.0006	0.18	0.003
HBGI	mean diff	-82.7	-4.0	-78.7
	p-value	0.003	1E-06	0.003
RI	mean diff	-123.3	-13.5	-109.8
	p-value	2E-05	0.03	0.0001

Table III: Comparison between the controller for all patients together. ( $\text{Ctrl}_1$  vs.  $\text{Ctrl}_2 \rightarrow \text{mean diff} = \text{Ctrl}_1 - \text{Ctrl}_2$ )

to 9pm. The average TIR is 47.7% for all groups together. With an average TIR of 26.9%, the performance of the BB controller for children is considerably lower than 55.9% and 59.7% TIR for adolescents and adults respectively. With this controller, the BG of children is 56.5% of the time below  $40 \frac{mg}{dl}$  which leads to an average LBGI of  $30.2 \frac{mg}{dl}$ . On average if adolescents are not in an euglycemic range they are more or less uniformly distributed in the other ranges with 11.5% TAR, 11.6% TBR, 8.7% THI, and 12.6% TLO. The adults are controlled best within these three groups. On average an adult spends 12.1% in a slight hyperglycemia, 16.7% in a slight hypoglycemia, 2.5% in severe hyperglycemia, and 8.3% in severe hypoglycemia. This leads to an LBGI of 5.5 and an overall RI of 7.1, which lie below the overall LBGI of 14.1 and RI of 17.0. On an individual level, the patient with the best BG levels achieves 97.2% TIR, 0% TAR, THI, and TLO and 2.8% TBR. Overall only five patients reach the goal of at least 70% TIR and at most 25% of TAR + THI. Of those five only 2 achieve a TBR + TLO of less than 4%.

#### B. PID controller

The performance of the PID controller is very much different from the performance of the BB controller. Looking at the plot 3b it is clear that the variability is very high. In general the mean is stable but on average above the desired range. Considering the days separately it is visible in figure 1b that the problem arises after 48h when the glycemia no longer is controlled. Overall, the average TIR is only 4.8% with the highest average of 6.7% for the adults. While children and adults are mainly in a hypoglycemic state with 98.1% and 78.2% TLO respectively, adults deviate to a severe hyperglycemic state for 81.9% of the day. This results in large variations in LBGI between the groups with 130.2, 33.2, and 0.9 for children, adolescents, and adults respectively. The patient with the best performance reaches a TIR of 22.01%. None of the patients reaches the goal of appropriate BG control.

### C. PPO controller

During the tuning, it was observed that changing  $\gamma$  and the learning rate for the PPO controller had no impact on the performance. Training an algorithm only with adults as the subject has proven to be moderately successful for adults and adolescents mainly. Training the algorithm only with children or adolescents respectively did not lead to improved performance of the controller for children or adolescents but was still considerably successful for adults. The statistical analysis has only been done for the discrete PPO controller. The mean curve of the discrete PPO is inside the target range for prolonged times and has excursions to hyperglycemia around 9am, 4pm, and 11pm. Excursions to hypoglycemia mainly occur around 4pm, 2pm and between 8pm and 9pm. Some of the patients probably had BG values above  $600 \frac{mg}{dl}$ , as can be seen by the plateaus for some of the gray curves in figure 3c. The PPO controller trained specifically for adults (figure 3d) has a low mean curve which often lies below the target range, with the most striking low between 7pm and 9pm where 95% of all patients have much lower BG than the target range. Similar results can be seen for the PPO controller trained for children (figure 3e). However, this controller keeps the BG in the target range for longer periods than the PPO controller of the adults. 53.5% of the time patients using the discrete PPO controller have a BG below  $50 \frac{mg}{dl}$ , 22.1% it is in the target range, 11.3% it lies above  $250 \frac{mg}{dl}$ , 7.7% slightly above and 5.4% slightly below the target range. The patient with the best BG control reaches a TIR of 61.95% and a TAR + THI of 21.8%. However, none of the patients reaches the goal of TIR above 70% or TBR + TLO below 4%.

### D. BB vs. PID

On average, the difference in TIR between BB and PID is 42.9% which is significant as demonstrated by the low p-value (1e-11). Conversely, for TAR and TBR the BB controller stays longer in these regions. Considering ranges where the BG is lower than the target range, the BB controller performs better as can be seen by the significant difference between the LBGI and the lower % TLO for the BB controller. For high BG levels, the BB controller performs better over all groups as well as for only adolescents or adults. However, for children, the BB controller does not perform better than the PID controller.

### E. BB vs. PPO

Comparing the 24h overlay plots for the BB and PPO controller (figures 3a and 3c) it is striking how similar the mean curve is. However, the SD is very different with a far greater SD for the PPO controller. Furthermore, the dip below the target range before lunch (around 2pm) and before dinner (between 7pm and 9pm) is larger for the PPO controller which suggests that with the PPO controller more severe hypoglycemic episodes occur around these times. This is supported by the fact that the average TLO is significantly higher with the PPO controller for adolescents, children, and over all patients. Although the average LBGI with the PPO is lower, the difference is not significant for all groups except

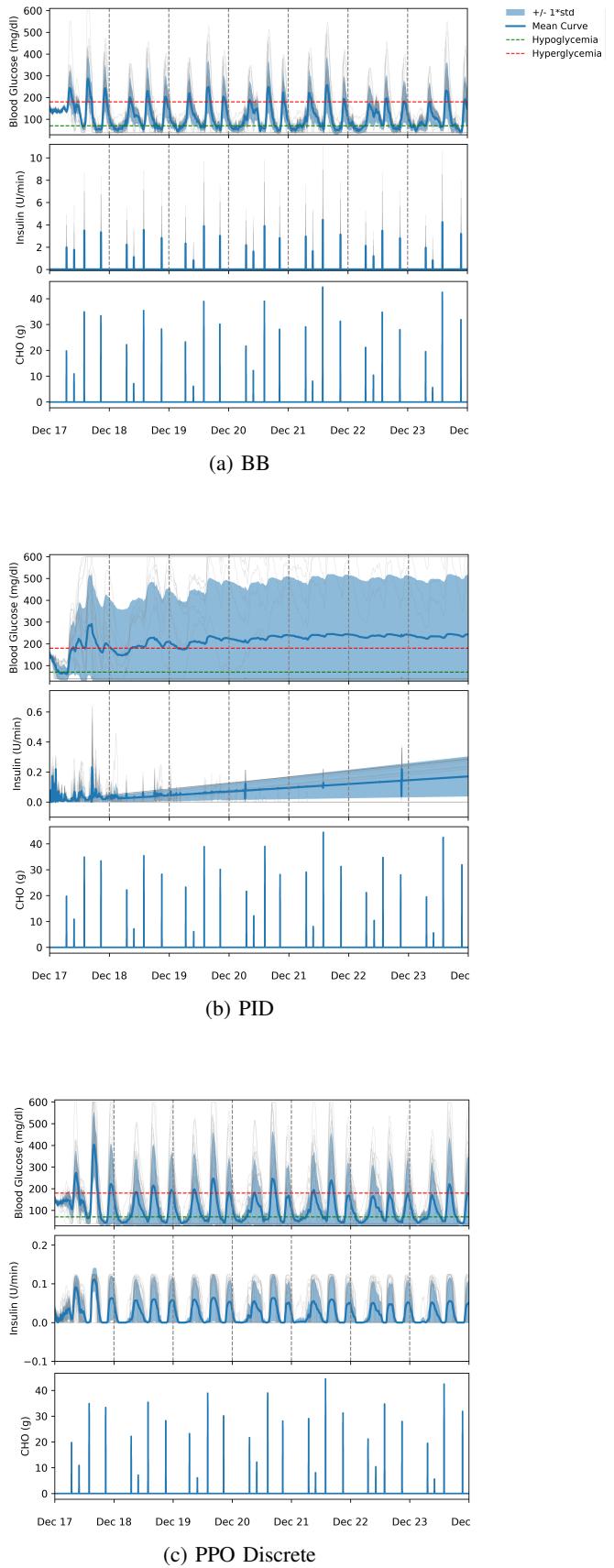
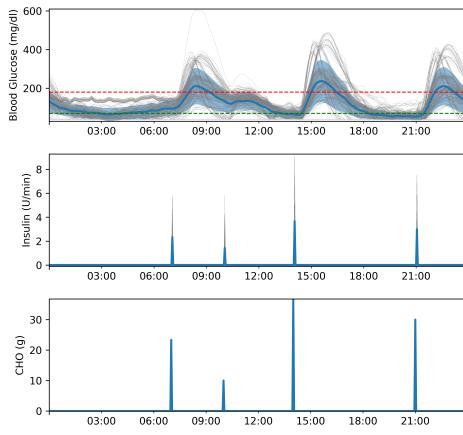
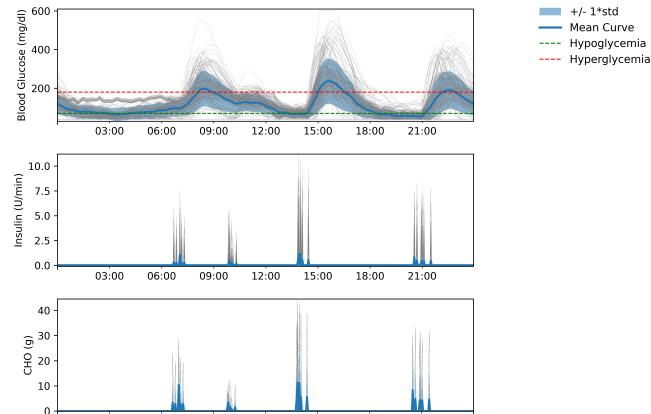


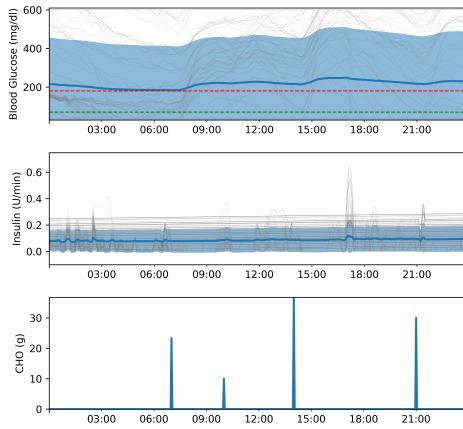
Fig. 1: Blood glucose profile simulated for 30 patients over 7 days with meal time and meal size variability



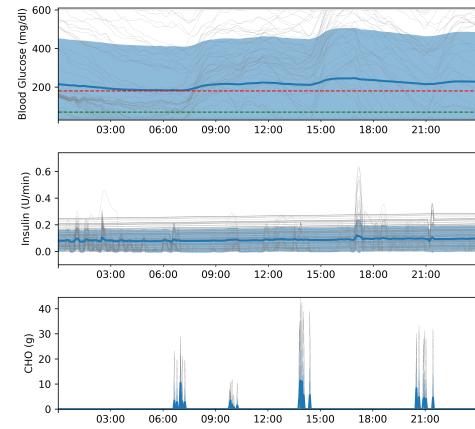
(a) BB



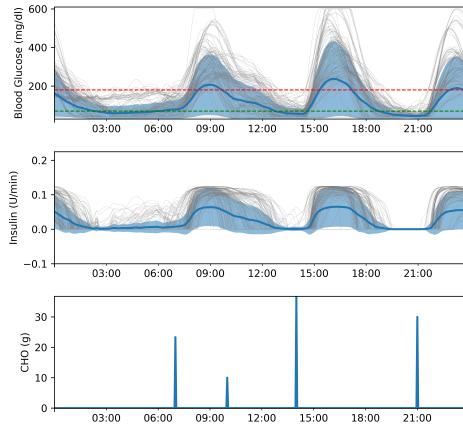
(b) BB



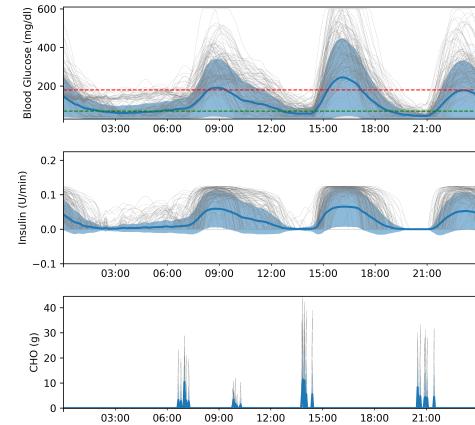
(c) PID



(d) PID



(e) PPO Discrete



(f) PPO Discrete

Fig. 2: Blood glucose profile, simulated for 30 patients over 7 days, shown as a 1 day overlay. a), c) and e) without variability. b), d) and f) with variability.

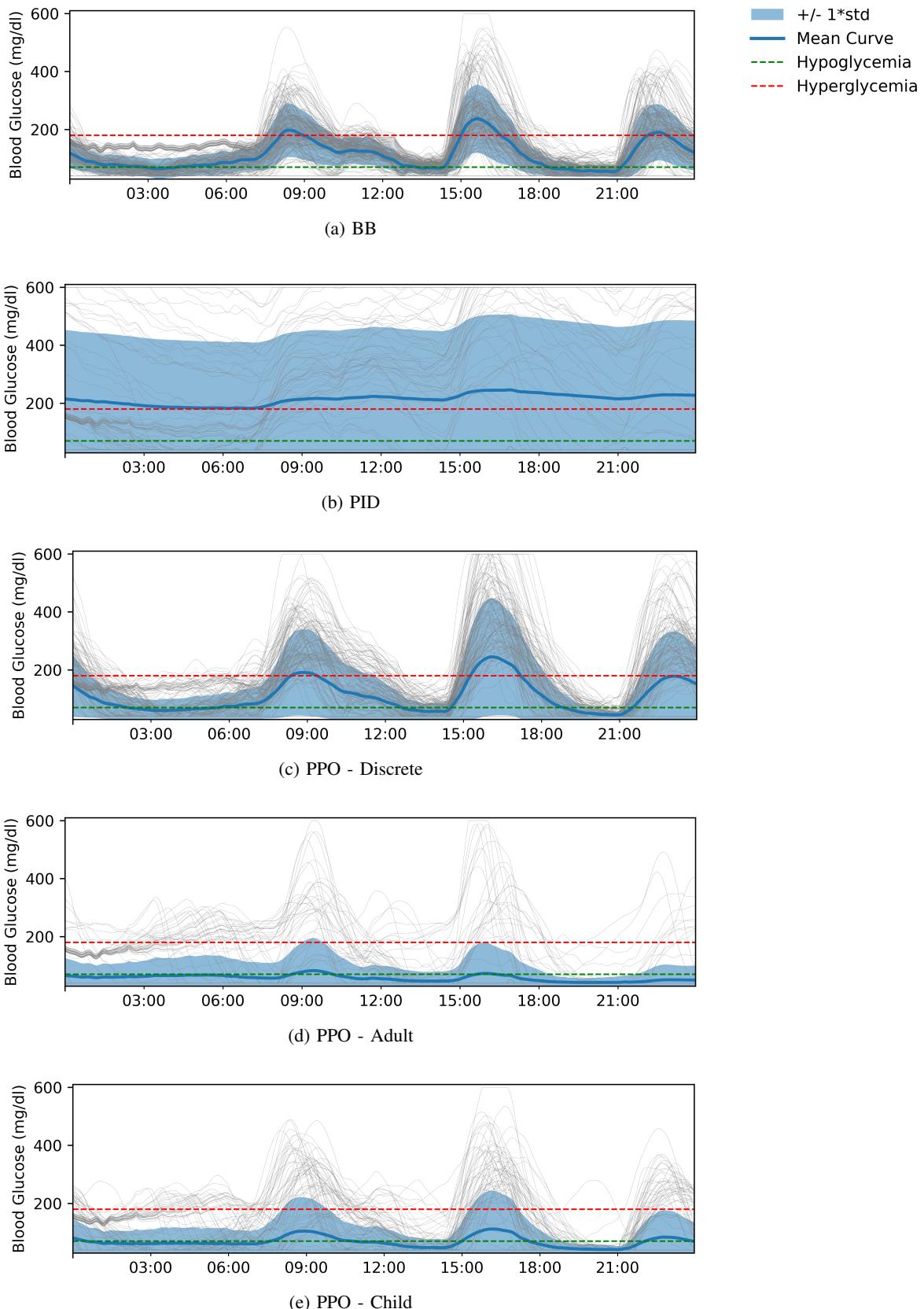


Fig. 3: Blood glucose profile, simulated for 30 patients over 7 days, shown as a 1 day overlay with meal time and meal size variability

adolescents where the difference is 12.4 with a p-value of 0.01. Comparing only the children of both controllers the difference is only significant considering TIR and TLO for all the other parameters the difference is not significant. However, when comparing the mean with the SD for the different parameters it can be seen that in many cases the order of magnitude is the same (e.g., TIR =  $7.6 \pm 8.9$ ).

#### F. PPO vs. PID

Looking at the graphs with the BG levels within 24h (figures 3c and 3b) the similarities are very limited. While the mean BG curve for the PPO controller shows a pattern where the eating times can be recognised, the PID controller shows a flat curve as the mean BG. The difference in TIR between PPO and PID is significant for all groups, the same can be stated for the RI. The PPO controller better manages THI for adults and adolescents, but not for children. On the other hand, the PPO controller has a better performance comparing TLO for children compared to PID. As for the other comparisons, the difference between TAR and TBR is not significant.

## IV. DISCUSSION

The mean BG curve of the BB controller and the PPO controller show similar patterns with the most striking difference being the SD. The PID controller fails to control the BG for more than 48h and therefore the mean BG curve is flat but most of the time above the target range. Although, the inter-patient variability is large in many cases some statistically relevant differences can be seen. Based on the % TIR, LBGI, and RI, the BB controller shows a higher performance than the PID controller in all individual groups as well as overall. The BB controller and PPO controller differ significantly considering only the % TIR. However, the LBGI is only significantly better in adolescents comparing these two controllers. The PPO controller performs better than the PID controller considering the % TIR and RI for all groups and overall. The LBGI is significantly better with the PPO versus PID controller for children and when averaging over all patients.

The high p-values comparing TAR and TBR (often 1 or very close to 1) might be due to the selection of the range for this parameter. While with the BB controller the peaks of the BG are close to this range, the TAR and TBR are higher than for PID and PPO controller, where the rise or the fall of the BG is steep, so that the TAR and TBR are rather low. The TAR for children where the PID controller is used is very low. However, this can be explained by the fact that for many children the insulin injected by the controller is too high and so they stay at a low BG level (in some cases the BG stayed at  $0 \frac{mg}{dl}$ , which is not a physiological BG value). Generally, when only considering the mean difference between two controllers the difference seems to be large. However, the t-test often shows no significant differences between the groups. This could be due to the large SD observed for many of the parameters, which means that the difference in means is mainly explained by chance and the probability that the two

groups are different is rather low. An interesting finding is that even if the PPO controller was only trained on children or adolescents the performance to control the BG of these groups was always low. Training the algorithm only with adults, adolescents, or children always had the best performance when simulating adults, even if the algorithm had not seen any adults previously. The explanation for this might be due to more than one reason. The effect might be eliminated by finding better parameters to train the algorithm. Another explanation might be that the glucodynamics and insulin sensitivity differs between adults and adolescents due to puberty [16]. However, the development of the insulin requirement over time and its factors are not yet fully understood.

Even with the BB controller, only very few reach the current goals for adequate BG control for pwT1D as the current state of the art suggests [3]. With the implemented controllers it seems to be especially hard to reach the goal of less than 4% hypoglycemia. The fact that the research done in this field is so vast, the task at hand is no easy feat. This short project proves that the PPO controller is a controller that can be used for BG control in pwT1D but needs to be tuned further. Furthermore, the PPO controller could surpass the PID controller's performance in this project. Nonetheless, the BB controller performs better in comparison to both the PID and PPO controller.

One limitation of this project is the large inter-patient variability. Although, the mean difference might suggest large differences in the performance of the controllers, many comparisons resulted in high p-values. Comparing only 10 patients per group is not sufficient to show whether there is a real difference between the different controllers for many parameters. The performance of the PID with these parameters is not sufficiently good to compete with the other controllers. However, if the tuning of the parameters is done systematically and considering more than 24h the performance of the PID could most probably be increased. For further research, the PID controller should be tested with different parameter sets and tuned more accurately. Since the LBGI judges all values below range, the implementation of the TBR could be changed to include the TLO to increase correlation with the LBGI. The same could be done for the TAR, so that this % also includes the THI. Another limitation is the BB controller as a baseline. Even though it works considerably well in some patients, the BB controller cannot reliably control the BG to meet the suggestions of adequate BG control [3].

In conclusion, it has been proven that the BB controller still is superior to both PID and PPO controller. However, these results are promising. With more hyperparameter tuning and improvements in the environments provided to the RL algorithm, the PPO controller might catch up with the performance of the BB controller.

## REFERENCES

- [1] P. Saeedi et al., “Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition,” *Diabetes Res Clin Pract*, vol. 157, p. 107843, Nov. 2019, doi: 10.1016/j.diabres.2019.107843.
- [2] “Diabetes.” <https://www.who.int/news-room/fact-sheets/detail/diabetes> (accessed Dec. 04, 2022).
- [3] T. Battelino et al., “Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range,” *Diabetes Care*, vol. 42, no. 8, pp. 1593–1603, Aug. 2019, doi: 10.2337/dc19-0028.
- [4] M. Tauschmann and R. Hovorka, “Technology in the management of type 1 diabetes mellitus — current status and future prospects,” *Nat Rev Endocrinol*, vol. 14, no. 8, Art. no. 8, Aug. 2018, doi: 10.1038/s41574-018-0044-y.
- [5] G. M. Steil, K. Rebrin, C. Darwin, F. Hariri, and M. F. Saad, “Feasibility of automating insulin delivery for the treatment of type 1 diabetes,” *Diabetes*, vol. 55, no. 12, pp. 3344–3350, Dec. 2006, doi: 10.2337/db06-0419.
- [6] T. Battelino, R. Nimri, K. Dovic, M. Phillip, and N. Bratina, “Prevention of Hypoglycemia With Predictive Low Glucose Insulin Suspension in Children With Type 1 Diabetes: A Randomized Controlled Trial,” *Diabetes Care*, vol. 40, no. 6, pp. 764–770, Jun. 2017, doi: 10.2337/dc16-2584.
- [7] R. Hovorka et al., “Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial,” *Lancet*, vol. 375, no. 9716, pp. 743–751, Feb. 2010, doi: 10.1016/S0140-6736(09)61998-X.
- [8] F. H. El-Khatib, S. J. Russell, D. M. Nathan, R. G. Sutherlin, and E. R. Damiano, “A bihormonal closed-loop artificial pancreas for type 1 diabetes,” *Sci Transl Med*, vol. 2, no. 27, p. 27ra27, Apr. 2010, doi: 10.1126/scitranslmed.3000619.
- [9] P. A. Bakhtiani, L. M. Zhao, J. El Youssef, J. R. Castle, and W. K. Ward, “A review of artificial pancreas technologies with an emphasis on bi-hormonal therapy,” *Diabetes Obes Metab*, vol. 15, no. 12, pp. 1065–1070, Dec. 2013, doi: 10.1111/dom.12107.
- [10] “jxx123/simglucose: A Type-1 Diabetes simulator implemented in Python for Reinforcement Learning purpose.” <https://github.com/jxx123/simglucose> (accessed Dec. 08, 2022).
- [11] J. Schulmann et al., “Proximal Policy Optimization Algorithms”. 2017, doi: 10.48550/ARXIV.1707.06347.
- [12] M. H. Yousuf, Autonomous-Insulin-Infusion-Controller. 2021. Accessed: Dec. 08, 2022. [Online]. Available: <https://github.com/MHamzaY/Autonomous-Insulin-Infusion-Controller>
- [13] openai/gym. OpenAI, 2022. Accessed: Dec. 08, 2022. [Online]. Available: <https://github.com/openai/gym>
- [14] R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [15] B. P. Kovatchev, D. J. Cox, L. A. Gonder-Frederick, and W. Clarke, “Symmetrization of the Blood Glucose Measurement Scale and Its Applications,” *DIABETES CARE*, vol. 20, no. 11, 1997.
- [16] G. H. Teló, C. E. Dougher, L. K. Volkening, M. L. Katz, and L. M. Laffel, “Predictors of changing insulin dose requirements and glycaemic control in children, adolescents and young adults with Type 1 diabetes,” *Diabet Med*, vol. 35, no. 10, pp. 1355–1363, Oct. 2018, doi: 10.1111/dme.13699.

## APPENDIX

		BB				PID				PPO discrete			
		all	chi	ado	adu	all	chi	ado	adu	all	chi	ado	adu
TIR	mean diff	-0.2	-0.2	0.3	-0.7	-0.4	-0.01	-0.8	-0.5	-0.8	-0.3	-0.6	-1.5
	p-value	0.23	0.71	0.73	0.02	0.20	0.47	0.04	0.01	1.0	0.8	1.0	0.9
TAR	mean diff	0.1	-0.1	0.4	-0.1	0.1	-0.03	0.3	-0.02	0.3	-0.1	0.1	0.9
	p-value	0.54	0.64	0.40	0.79	0.18	0.21	0.17	0.92	0.9	0.2	0.7	0.8
TBR	mean diff	0.8	0.2	1.4	0.9	-0.1	-0.01	-0.3	0.04	0.2	-0.3	-0.1	1.0
	p-value	0.02	0.43	0.002	0.0001	0.33	0.02	0.13	0.23	0.8	0.2	0.2	1.0
THI	mean diff	0.6	0.5	1.2	0.1	1.8	0.1	4.8	0.6	0.5	-0.1	0.6	1.1
	p-value	0.25	0.06	0.14	0.70	0.29	0.08	0.12	0.01	0.9	0.2	0.9	0.9
TLO	mean diff	-1.3	-0.4	-3.3	-0.2	-1.4	-0.01	-4.0	-0.1	-0.2	0.8	0.0	-1.4
	p-value	0.18	0.67	0.16	0.33	0.28	0.36	0.14	0.04	0.4	1.0	0.5	0.3
LBDGI	mean diff	0.7	4.1	-1.6	-0.3	-0.5	-0.2	-1.1	-0.03	-5.4	-14.4	-0.8	-0.9
	p-value	0.65	0.08	0.34	0.01	0.85	0.41	0.49	0.67	0.03	0.1	0.2	0.2
HBGI	mean diff	-0.1	-0.02	-0.3	-0.1	3.8	-0.1	11.7	-0.2	-0.3	-0.1	-0.1	-0.6
	p-value	0.08	0.91	0.01	0.0001	0.46	0.21	0.11	0.77	0.1	0.4	0.4	0.0003
RI	mean diff	0.6	4.1	-1.9	-0.4	3.4	-0.3	10.6	-0.2	-5.6	-14.5	-0.9	-1.5
	p-value	0.59	0.08	0.28	0.001	0.42	0.30	0.08	0.74	0.03	0.1	0.2	0.1

Table IV: Comparison between the scenario with variation and without. Paired t-test with  $\alpha = 0.95$ . Results in green show significant differences between the means ( $p < 0.05$ )

		BB vs. PID no variation				BB vs. PID variation				BB vs. PPO no variation				BB vs. PPO variation				PPO vs. PID no variation				PPO vs. PID variation				
		all	chi	ado	adu	all	chi	ado	adu	all	chi	ado	adu	all	chi	ado	adu	all	chi	ado	adu	all	chi	ado	adu	
TIR	mean 1	47.5	26.9	55.9	59.7	47.7	27.1	55.6	60.4	47.5	26.9	55.9	59.7	47.7	27.1	55.6	60.4	21.3	7.6	20.2	35.9	22.1	7.9	20.8	37.5	
	mean 2	4.3	1.5	5.4	6.2	4.8	1.5	6.1	6.7	21.3	7.6	20.2	35.9	22.1	7.9	20.8	37.5	4.3	1.5	5.4	6.2	4.8	1.5	6.1	6.7	
	SD 1	23.6	20.3	23.6	10.1	24.2	20.7	25.0	10.4	23.6	20.3	23.6	10.1	24.2	20.7	25.0	10.4	18.0	8.9	20.0	10.8	18.5	8.3	19.8	12.2	
	SD 2	3.7	0.7	4.4	2.9	4.6	0.7	6.3	3.1	18.0	8.9	20.0	10.8	18.5	8.3	19.8	12.2	3.7	0.7	4.4	2.9	4.6	0.7	6.3	3.1	
	p-value	7E-12	0.002	3E-05	6E-08	1E-11	0.002	3E-11	5E-12	2E-11	0.002	4E-05	5E-08	2E-09	0.002	3E-05	0.001	7E-07	0.03	0.01	7E-07	6E-07	0.02	0.01	2E-06	
TAR	mean 1	9.3	3.8	11.9	12.1	9.2	3.9	11.5	12.1	9.3	3.8	11.9	12.1	9.2	3.9	11.5	12.1	8.0	1.3	5.2	17.4	7.7	1.3	5.1	16.5	
	mean 2	1.4	0.1	1.7	2.6	1.4	0.2	1.3	2.6	8.0	1.3	5.2	17.4	7.7	1.3	5.1	16.5	1.4	0.1	1.7	2.6	1.4	0.2	1.3	2.6	
	SD 1	6.7	4.2	6.6	5.8	6.5	4.1	6.6	5.5	6.7	4.2	6.6	5.8	6.5	4.1	6.6	5.5	8.9	2.2	5.5	7.9	8.0	2.1	5.2	6.0	
	SD 2	2.1	0.3	2.4	2.1	1.9	0.3	1.5	2.2	8.9	2.2	5.5	7.9	8.0	2.1	5.2	6.0	2.1	0.3	2.4	2.1	1.9	0.3	1.5	2.2	
	p-value	1.0	0.99	1.0	1.0	1.0	0.99	1.0	1.0	1.0	0.90	1.0	0.87	0.75	0.92	0.96	0.07	1.00	0.93	0.99	1.00	1.00	0.94	0.99	1.00	
TBR	mean 1	13.0	8.6	12.9	17.6	12.2	8.4	11.6	16.7	13.0	8.6	12.9	17.6	12.2	8.4	11.6	16.7	5.6	1.5	3.6	11.7	5.4	1.8	3.7	10.7	
	mean 2	0.5	0.2	0.9	0.4	0.6	0.3	1.2	0.3	5.6	1.5	3.6	11.7	5.4	1.8	3.7	10.7	0.5	0.2	0.9	0.4	0.6	0.3	1.2	0.3	
	SD 1	8.0	7.8	7.3	6.7	7.6	7.6	6.9	6.4	8.0	7.8	7.3	6.7	7.6	7.6	6.9	6.4	5.8	2.7	3.8	4.8	5.5	3.5	4.0	4.4	
	SD 2	0.8	0.1	1.0	0.7	1.2	0.2	2.0	0.6	5.8	2.7	3.8	4.8	5.5	3.5	4.0	4.4	0.8	0.1	1.0	0.7	1.2	0.2	2.0	0.6	
	p-value	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.99	1.00	0.91	0.99	1.00	0.89	0.99	1.00		
THI	mean 1	5.7	4.7	9.9	2.6	5.1	4.2	8.7	2.5	5.7	4.7	9.9	2.6	5.1	4.2	8.7	2.5	11.9	3.0	12.4	20.3	11.3	3.1	11.7	19.2	
	mean 2	33.3	0.1	18.0	81.9	31.5	0.1	13.2	81.3	11.9	3.0	12.4	20.3	11.3	3.1	11.7	19.2	33.3	0.1	18.0	81.9	31.5	0.1	13.2	81.3	
	SD 1	7.5	6.4	9.8	4.0	6.6	5.6	5.6	8.7	3.3	7.5	6.4	9.8	4.0	6.6	5.6	8.7	3.3	12.0	5.5	13.3	9.6	11.1	5.3	12.6	8.0
	SD 2	43.3	0.3	33.6	28.3	43.0	0.1	30.2	28.2	12.0	5.5	13.3	9.6	11.1	5.3	12.6	8.0	43.3	0.3	33.6	28.3	43.0	0.1	30.2	28.2	
	p-value	0.001	0.98	0.26	4E-06	0.001	0.98	0.002	0.002	0.98	0.93	0.98	0.27	0.01	0.72	0.26	5E-05	0.002	0.93	0.30	0.00001	0.002	0.95	0.30	1E-05	
TLO	mean 1	24.5	56.0	9.4	8.1	25.8	56.5	12.6	8.3	24.5	56.0	9.4	8.1	25.8	56.5	12.6	8.3	53.3	86.7	58.5	14.7	53.5	85.9	58.6	16.1	
	mean 2	60.4	98.1	74.2	9.0	61.8	98.1	78.2	9.1	53.3	86.7	58.5	14.7	53.5	85.9	58.6	16.1	60.4	98.1	74.2	9.0	61.8	98.1	78.2	9.1	
	SD 1	29.8	31.9	11.1	7.8	30.8	31.8	19.7	7.5	29.8	31.9	11.1	7.8	30.8	31.8	19.7	7.5	38.4	19.2	35.5	13.7	37.7	19.1	35.3	15.3	
	SD 2	46.0	1.3	37.2	26.8	45.2	1.3	31.7	26.8	38.4	19.2	35.5	13.7	37.7	19.1	35.3	15.3	46.0	1.3	37.2	26.8	45.2	1.3	31.7	26.8	
	p-value	2E-05	0.001	0.00001	0.46	2E-05	0.001	2E-05	6E-06	1E-12	0.002	1E-09	3E-06	2E-05	0.004	0.001	0.13	0.10	0.05	0.08	0.71	0.10	0.04	0.08	0.74	
LBGI	mean 1	14.9	34.3	5.1	5.2	14.1	30.2	6.7	5.5	14.9	34.3	5.1	5.2	14.1	30.2	6.7	5.5	18.3	29.3	18.3	7.3	23.6	43.7	19.1	8.2	
	mean 2	54.3	130.0	32.1	0.9	54.8	130.2	33.2	0.9	18.3	29.3	18.3	7.3	23.6	43.7	19.1	8.2	54.3	130.0	32.1	0.9	54.8	130.2	33.2	0.9	
	SD 1	22.0	30.0	4.7	3.5	19.2	24.5	9.6	3.9	22.0	30.0	4.7	3.5	19.2	24.5	9.6	3.9	13.4	11.4	11.8	6.7	23.5	28.4	13.5	7.8	
	SD 2	70.8	71.7	30.4	2.1	70.6	71.4	29.8	2.1	13.4	11.4	11.8	6.7	23.5	28.4	13.5	7.8	70.8	71.7	30.4	2.1	70.6	71.4	29.8	2.1	
	p-value	0.001	0.001	0.01	1.0	0.00006	0.00006	0.001	0.0003	0.009	0.54	0.004	0.002	0.18	0.70	0.01	0.21	0.002	0.001	0.06	0.09	0.003	0.001	0.08	0.99	
HBGI	mean 1	2.7	2.9	3.7	1.5	2.8	2.9	4.0	1.6	2.7	2.9	3.7	1.5	2.8	2.9	4.0	1.6	6.6	2.9	9.5	7.4	6.8	3.0	9.6	8.0	
	mean 2	89.3	0.1	53.1	214.9	85.5	0.2	41.4	215.1	6.6	2.9	9.5	7.4	6.8	3.0	9.6	8.0	89.3	0.1	53.1	214.9	85.5	0.2	41.4	215.1	
	SD 1	2.9	3.9	2.9	0.7	3.0	3.9	3.1	0.8	2.9	3.9	2.9	0.7	3.0	3.9	3.1	0.8	4.6	3.0	5.4	2.2	4.7	2.7	5.8	2.3	
	SD 2	153.1	0.2	101.9	193.2	152.9	0.3	98.1	191.9	4.6	3.0	5.4	2.2	4.7	2.7	5.8	2.3	153.1	0.2	101.9	193.2	152.9	0.3	98.1	191.9	
	p-value	0.002	0.97	0.08	0.003	0.97	0.003	0.004	1E-10	0.006	4E-05	8E-08	1E-06	0.47	0.0001	1E-06	0.003	0.99	0.11	0.004	0.003	1.00	0.12	0.004	0.004	
RI	mean 1	17.6	37.2	8.8	6.7	17.0	33.1	10.7	7.1	17.6	37.2	8.8	6.7	17.0	33.1	10.7	7.1	24.9	32.2	27.7	14.7	30.5	46.7	28.6	16.2	
	mean 2	143.7	130.1	85.2	215.8	140.3	130.4	74.5	216.0	24.9	32.2	27.7	14.7	30.5	46.7	28.6	16.2	143.7	130.1	85.2	215.8	140.3	130.4	74.5	216.0	
	SD 1	22.0	29.1	6.9	3.9	19.3	23.6	12.0	4.4	22.0	29.1	6.9	3.9	19.3	23.6	12.0	4.4	12.8	11.4	12.3	7.6	21.9	27.3	14.0	9.0	
	SD 2	136.5	71.6	89.5	192.7	137.4	71.2	88.3	191.5	12.8	11.4	12.3	7.6	21.9	27.3	14.0	9.0	136.5	71.6	89.5	192.7	137.4	71.2	88.3	191.5	
	p-value	2E-05	0.002	0.01	0.004	2E-05	0.001	3E-05	2E-05	0.002	0.49	0.002	0.0004	0.03	0.70	0.0003	0.01	3E-05	0.001	0.04	0.004	0.0001	0.002	0.05	0.004	

Table V: Comparison of difference of mean for the different controller. Paired t-test testing for increased TIR, decreased TAR, THI, TBR, TLO, LGBI, HBGI, and RI. Significance level ( $\alpha$ ) = 0.95. P-values marked green are  $< 0.05$  and the difference between the compared variables considered significant. (Ctrl<sub>1</sub> vs. Ctrl<sub>2</sub> → mean 1 and SD 1 of Ctrl<sub>1</sub> or mean 2 and SD 2 of Ctrl<sub>2</sub>)