ESCADA: Efficient Safety and Context Aware Dose Recommendation for Precision Medicine

Machine Learning for Healthcare and Diabetes Workshop

Ilker Demirel

Bilkent University
Department of Electrical and Electronics Engineering
Cognitive Systems, Bandits, and Optimization Research Group

January 19, 2022

CYBORG

Multi Armed Bandits (MAB)

Figure 1: MAB Framework

- a simple "online" learning problem
- **m** arms to choose one from at each round
- arms have **unknown** reward yielding patterns

Multi Armed Bandits (MAB)

Figure 1: MAB Framework

- a simple "online" learning problem
- **m** arms to choose one from at each round
- arms have **unknown** reward yielding patterns
- how should a **learner** pick arms over rounds to maximize its expected earnings in the long run?

Exploration vs. Exploitation

- **exploitation:** choose the arm that looks best based on past observations
  - you may stuck at a "local" minimum

Exploration vs. Exploitation

- **exploitation:** choose the arm that looks best based on past observations
  - you may stuck at a "local" minimum

	arm 1	arm 2	arm 3	arm 4	arm 5
p	0.1	0.2	0.3	**0.8**	**0.9**
$\hat{p}$	0	0.25	0.25	**0.75**	**0.5**

Table 1: **True** (p) and **Observational** ($\hat{p}$) reward yielding probabilities for Bernoulli arm rewards (after 4 trials for each arm)

Exploration vs. Exploitation

- **exploitation:** choose the arm that looks best based on past observations
  - you may stuck at a "local" minimum

	arm 1	arm 2	arm 3	arm 4	arm 5
p	0.1	0.2	0.3	**0.8**	**0.9**
$\hat{p}$	0	0.25	0.25	**0.75**	**0.5**

Table 1: **True** (p) and **Observational** ($\hat{p}$) reward yielding probabilities for Bernoulli arm rewards (after 4 trials for each arm)

- **exploration:** try different arms to gather as much information as possible
  - why keep playing an arm that is almost certainly suboptimal?

Gaussian Processes

- learn **non-parametric** functions (for instance, $y = ax + b$ is parametric)
- analytically tractable
- confidence intervals

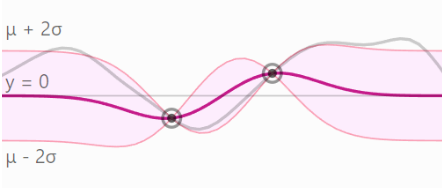


Figure 2: Gaussian Process

Gaussian Processes

- learn **non-parametric** functions (for instance, $y = ax + b$ is parametric)
- analytically tractable
- confidence intervals

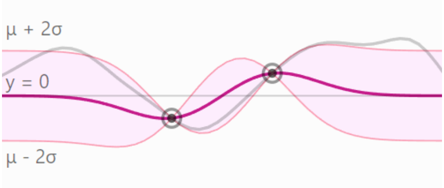


Figure 2: Gaussian Process

- **can model the correlations between arm outcomes**

Application

- bolus-insulin dose recommendation for **type-1 diabetes**
- **blood glucose control**
- **hypo-/hyper-glycemia** events: dangerously low and high blood glucose levels

Application

- bolus-insulin dose recommendation for **type-1 diabetes**
- **blood glucose control**
- **hypo-/hyper-glycemia** events: dangerously low and high blood glucose levels

can we design...

Application

- bolus-insulin dose recommendation for **type-1 diabetes**
- **blood glucose control**
- **hypo-/hyper-glycemia** events: dangerously low and high blood glucose levels

can we design...

- a personalized, data-driven, and safe treatment optimization strategy,

Application

- bolus-insulin dose recommendation for **type-1 diabetes**
- **blood glucose control**
- **hypo-/hyper-glycemia** events: dangerously low and high blood glucose levels

can we design...

- a personalized, data-driven, and safe treatment optimization strategy,
- without calculating parameters such as *insulin to carbohydrate ratio* (ICR), or *correction factor* (CF)

Can we use MABs?

- a **MAB** model for each patient

Can we use MABs?

- a **MAB** model for each patient
- **arms:** bolus-insulin dose recommendations

Can we use MABs?

- a **MAB** model for each patient
- **arms:** bolus-insulin dose recommendations
- **reward:** proximity of postprandial blood glucose level (BGL) to the target BGL

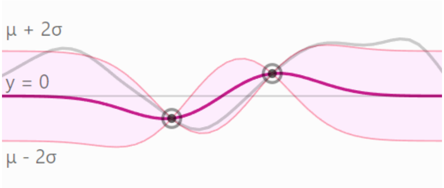


Figure 3: use a GP to learn a mapping from bolus-insulin doses to postprandial BGL

Mathematical Summary

Goal

learn recommending doses which trigger the closest postprandial BGLs to the target BGL

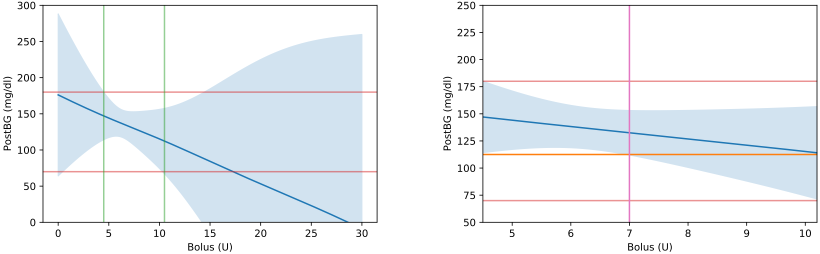
$$\text{minimize} \quad R_N = \sum_{n=1}^{N} \left| f\left(z_n, d_n\right) - T \right|$$

Constraint

Safe Exploration: never recommend doses which lead to **hypo-/hyper-glycemia** events

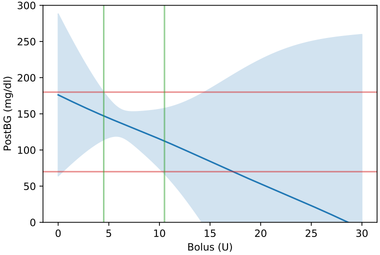
$$\text{subject to} \quad T_{\min} \leq f\left(z_n, d_n\right) \leq T_{\max}$$

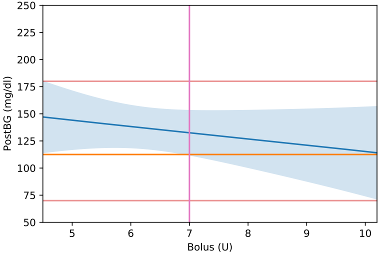
ESCADA Algorithm



- red horizontal lines are the hyper-/hypo-glycemia boundaries, and the orange horizontal line is the target BGL

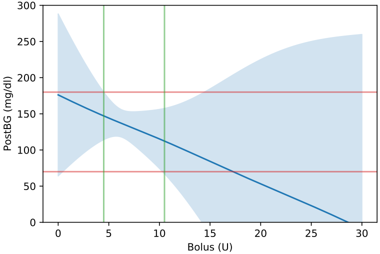
ESCADA Algorithm

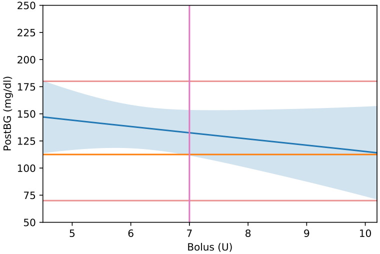




- red horizontal lines are the hyper-/hypo-glycemia boundaries, and the orange horizontal line is the target BGL
- between vertical green lines are the safe doses

ESCADA Algorithm

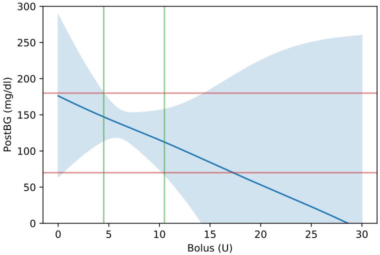


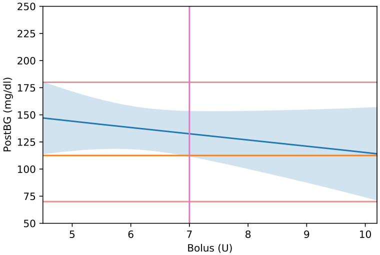


- red horizontal lines are the hyper-/hypo-glycemia boundaries, and the orange horizontal line is the target BGL
- between vertical green lines are the safe doses

- observe the meal event context (z_n)
- **context:** fasting BGL, carbohydrate content, exercise after meal etc.

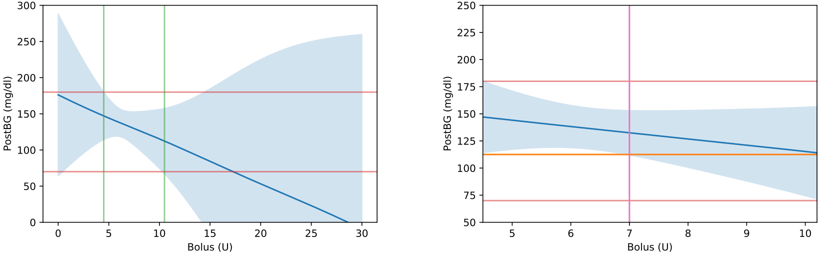
ESCADA Algorithm





- red horizontal lines are the hyper-/hypo-glycemia boundaries, and the orange horizontal line is the target BGL
- between vertical green lines are the safe doses
- observe the meal event context (z_n)
- **context:** fasting BGL, carbohydrate content, exercise after meal etc.
- identify the safe doses whose **confidence intervals** contain the target BGL
- recommend the dose with closest **mean response** to the target BGL

ESCADA Algorithm



- red horizontal lines are the hyper-/hypo-glycemia boundaries, and the orange horizontal line is the target BGL
- between vertical green lines are the safe doses
- observe the meal event context (z_n)
- **context:** fasting BGL, carbohydrate content, exercise after meal etc.
- identify the safe doses whose **confidence intervals** contain the target BGL
- recommend the dose with closest **mean response** to the target BGL
- if there is no such dose, recommend the most uncertain dose (**exploration**)

Recap

- multi-armed bandits (MAB)

Recap

- multi-armed bandits (MAB)
- exploration vs. exploitation

Recap

- multi-armed bandits (MAB)
- exploration vs. exploitation
- bolus-insulin dose recommendation as a MAB problem

Recap

- multi-armed bandits (MAB)
- exploration vs. exploitation
- bolus-insulin dose recommendation as a MAB problem
- ESCADA: a **personalized** dose recommendation algorithm with **safety** and **performance** guarantees

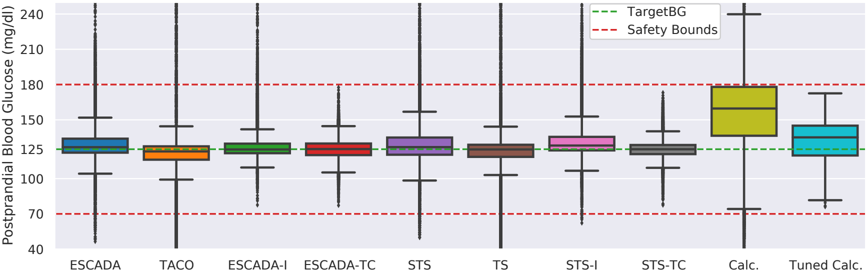
The Simulator

- 30 different virtual patients: 10 children, 10 adolescents, 10 adults

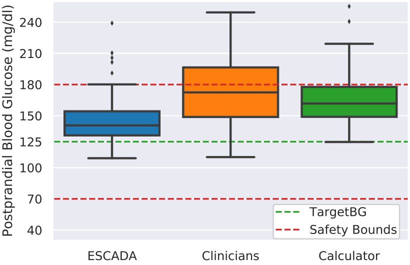
The Simulator

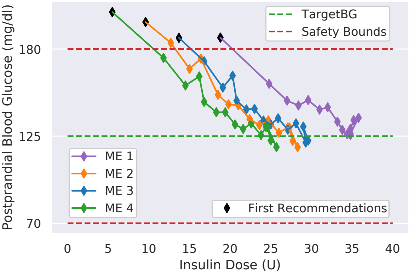
- 30 different virtual patients: 10 children, 10 adolescents, 10 adults
- BGL response to bolus-insulin dose is modeled via complex **differential equations**
- **U.S. FDA** verified for closed loop hormone controller design

Experiments with Simulator



Experiments with the Simulator





CYBORG

Clinician Screens

hasta 1

	carb	bolus	min	tmbg	fastbg	postbg
25	30.0	6	480	150	136.0	136.0
27	30.0	6	480	150	171.0	115.0

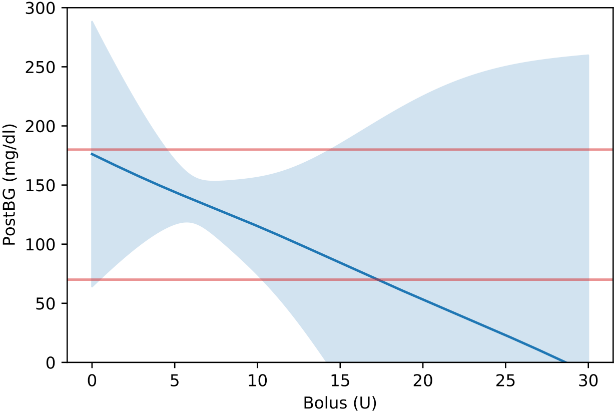
hasta 2

	carb	bolus	min	tmbg	fastbg	postbg
10	55	6	540	150	119.0	127.0
11	55	6	540	150	90.0	170.0

Clinician Screens

```
Patient: hasta 9, Time: 7:20:00, Carb: 30 g, FastBG: 150 mg/dl, TMBG: 160 mins

Dose: 17.2 U, 95% CI: (0, 229) mg/dl, Safe w.p.: 0.63 (70, 180) mg/dl, Hyper w.p.: 0.13 (>180 mg/dl), Hypo w.p.: 0.24 (<70 mg/dl)

Safest Available Doses:

Dose: 10.3 U, 95% CI: (163, 181) mg/dl, Safe w.p.: 0.95 (70, 180) mg/dl, Hyper w.p.: 0.05 (>180 mg/dl), Hypo w.p.: 0.00 (<70 mg/dl)

Queried Doses:

Dose: 12.0 U, 95% CI: (119, 194) mg/dl, Safe w.p.: 0.89 (70, 180) mg/dl, Hyper w.p.: 0.11 (>180 mg/dl), Hypo w.p.: 0.00 (<70 mg/dl)
Dose: 13.0 U, 95% CI: (93, 202) mg/dl, Safe w.p.: 0.87 (70, 180) mg/dl, Hyper w.p.: 0.13 (>180 mg/dl), Hypo w.p.: 0.00 (<70 mg/dl)
Dose: 14.0 U, 95% CI: (68, 210) mg/dl, Safe w.p.: 0.84 (70, 180) mg/dl, Hyper w.p.: 0.13 (>180 mg/dl), Hypo w.p.: 0.03 (<70 mg/dl)
```

Clinician Screens



CYBORG