Algorithms Among Us: Artificial Intelligence in Closed-Loop Clinical Treatment Applications

Chirath Hettiarachchi

Research Fellow,
School of Computing, College of Systems & Society
Australian National University

November 27, 2025

Acknowledgement of Country

Disclaimer

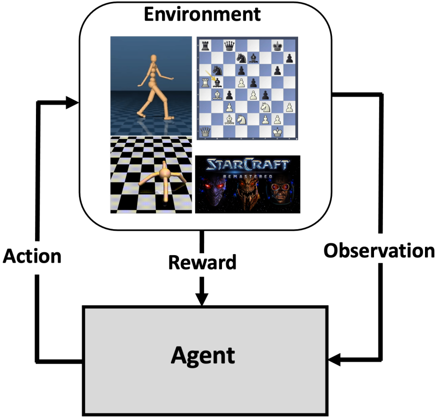
- The algorithms I present today are **not among us yet!**
  (its a good thing, we have more time to prepare...).
- The talk is limited to our research on a subset of AI algorithms called **RL**.
  (NOT Generative AI, LLM's, or standard classification/regression models).
- **Beyond the algorithm**: Not a technical talk, sharing my experience / thoughts on the future of AI in medicine. *(therefore, I'm not an expert and love to learn from you).*

■ A talk prepared for the Political and Environmental Psychology & Social Science (PEPSS) Seminar, hosted by the SMP Seminar series. Theme: **Algorithms Among Us: AI's Expanding Role in Public Life**, with sub-themes ideally centered around democracy and/or health (Democracy in the Age of AI: Threats, Promises, & Policy Gaps; AI in Public/Mental Health: Efficiency, Ethics, and Systemic Risk; AI in Health Systems: Trust, Risk, & the Psychology of Care).

Overview

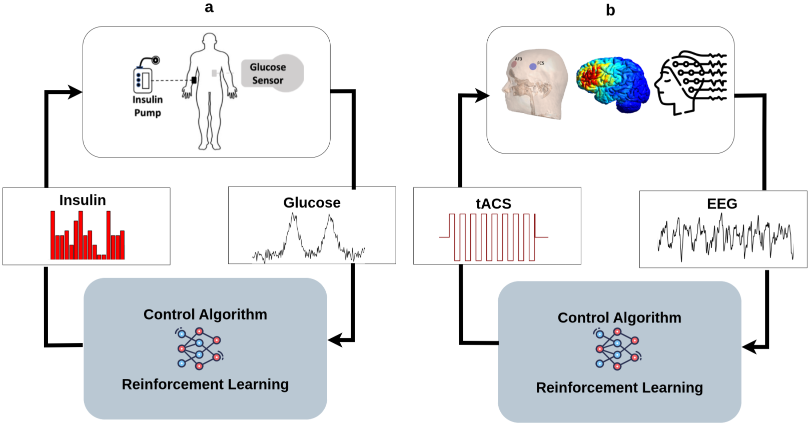
1. Background: Reinforcement Learning (RL) & closed-loop clinical treatment
2. Shaping a future together: Clinicians, lived experience, and AI
3. Building trust through education: It's easier to trust when you understand...
4. Ensuring fairness and equity
5. Engineering solutions for consent, privacy, and security
6. Accountability and need for regulatory frameworks
7. Conclusion

Content



The “Law of Effect” in psychology introduced learning by trial and error, which described the effect of reinforcing events based on the tendency to select actions [Thorndike, 1911].

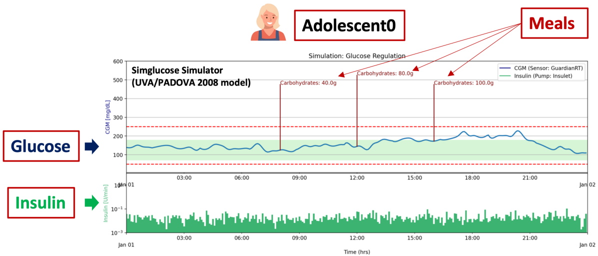
Closed-loop clinical treatment applications



(a) Glucose regulation in type 1 diabetes — insulin is administered to manage glucose and **(b)** Transcranial stimulation for depression — EEG guided transcranial alternating current stimulation (tACS) to modulate abnormal neural oscillations.

Example RL treatment decisions

a

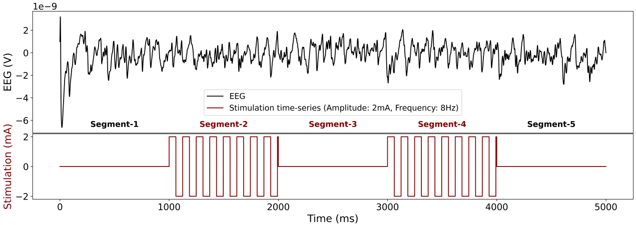


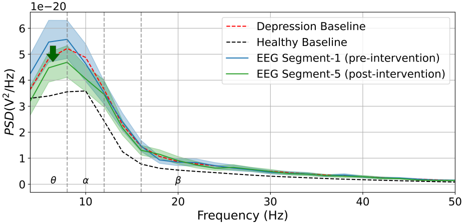
Control Algorithm (e.g., *our G2P2C algorithm*)

Table. Clinical performance metrics.

	Metric	Value
0	Time Above Range (TAR) - Level2 (>250 mg/dL)	0.0%
1	Time Above Range (TAR) - Level 1 (180 - 250 mg/dL)	14.93%
2	Time In Range (TIR) (70 - 180 mg/dL)	85.07%
3	Time Below Range (TIR) - Level 1 (54 - 70 mg/dL)	0.0%
4	Time Below Range (TIR) - Level 2 (<54 mg/dL)	0.0%
5	Risk Index (RI)	4.21
6	Low Blood Glucose Index (LBGI)	0.44
7	High Blood Glucose Index (HBGI)	3.77

b





Content

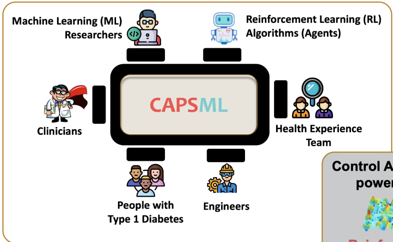
“Our Health in Our Hands” (OHIOH) [ANU 2017 Grand Challenge]

“We [lived experience experts] were privileged to enough to participate in a workshop [...] to brainstorm and evaluate the important traits of our ideal management device. [...] I found it to be one of the most rewarding and engaging exercises I’ve had the opportunity to work on.

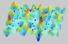
I really felt [...] the session allowed us to ***consider & appreciate ideas from a variety of perspectives, including sociologists, diabetes educators, social workers, as well as our own lived experience members.*** *[...].*

I can honestly say that my ideas and the ideas of my fellow lived experience researchers were properly and thoroughly considered alongside the advice from healthcare professionals which is something I’m really honoured by.”
- Young person with T1D [`https://youtu.be/1B1dp8GMZps`*]*

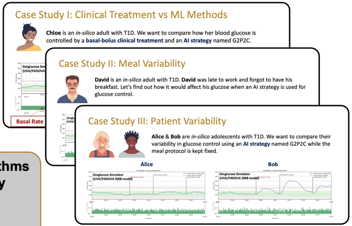
CAPSML: Co-creation & synergies when people and AI work together



Control Algorithms powered by



Reinforcement Learning



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

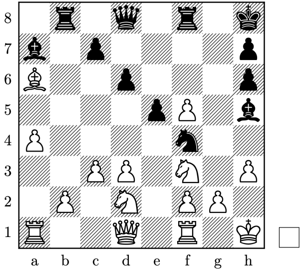
Learn about your meals
How much carbs are in a **McDonalds Cheese Burger**?
How does it affect glucose?

Desborough, et al. "A Framework for Involving Coproduction Partners in Research About Young People with Type 1 Diabetes", Health Expectations, 2022.

Hettiarachchi, et al. "CAPSML: Bridging the Gap Between Clinicians, Lived Experience Experts, and Artificial Intelligence Systems for Glucose Regulation in Type 1 Diabetes", 34^{th} Medical Informatics Europe (MIE), 2023.

Kasparov's Law:

- "**Weak human + machine + better process** was superior to a **strong computer** alone, and more remarkably, superior to a **strong human + machine + inferior process**."



Stockfish 8 vs AlphaZero
London 2018(12) (0-1)

Kasparov, G., 2017. Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins. Hachette, UK. 115.

■ AlphaZero sacrifices three pawns and the kingside structure to open lines all over the board and ultimately wins the game (Game Changer 2019, Matthew Sadler & Natasha Regan).

AI in Medicine: Shaping a future together (improving human capabilities)

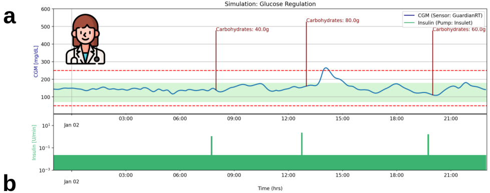
- Train **strong humans** using **AI**.
  We must improve our knowledge & skills by integrating AI tools, new insights they produce into curricula, thereby up-skilling and **strengthening core expertise**.

- In our research, we are trying to **discover new knowledge** using **AI** helpful for clinical treatment. This new knowledge will ideally influence clinicians in the short-term (**human-in-the-loop**).

  Once the technology matures, and we are stronger with a good understanding, we could focus on full automation, then move to the next challenging problem...

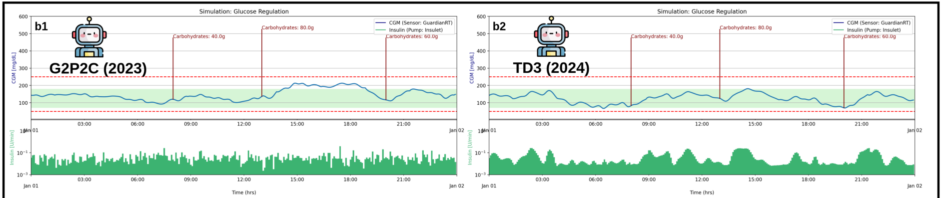
Content

It’s easier to trust when you understand...



Clinical treatment strategy (basal-bolus), easy to understand.

Requires manual user input on meal announcement and carbohydrate estimation adding significant cognitive burden!

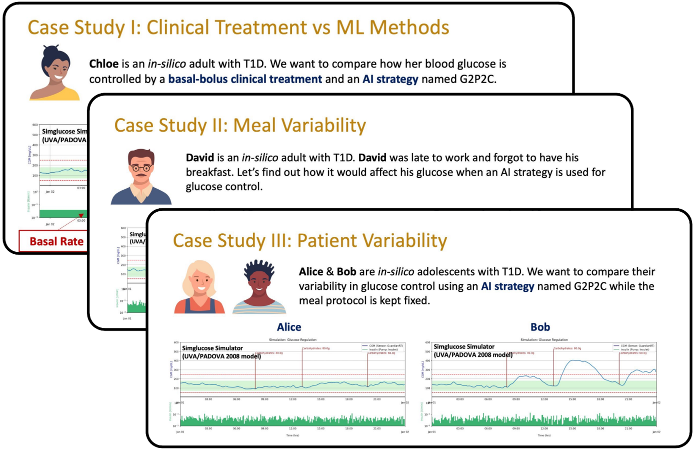


Fully automated RL-based insulin delivery systems. However, harder to interpret (b1), focussing on improving explainability (b2).

Hettiarachchi, et al.“G2P2C—A modular reinforcement learning algorithm for glucose control by glucose prediction & planning in T1D”, Biomedical Signal Processing & Control, 90, p.105839, 2023.

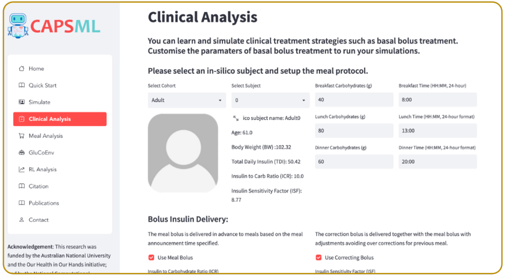
Timms, David, et al. “Comparing Deterministic and Stochastic Reinforcement Learning for Glucose Regulation in Type 1 Diabetes”, 20th World Congress on Medical and Health Informatics: MedInfo 2025.

Building tools for education: Hands-on with closed-loop AI systems

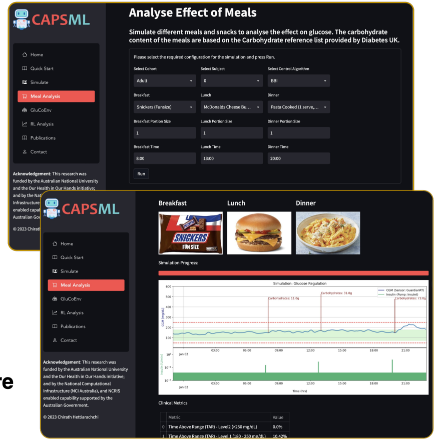


Develop your own case studies to learn / stress-test our closed-loop AI systems, through simulations and analysis.

Building tools for education: Requests by clinicians & people with T1D



A **user-friendly, accessible tool** for analysing existing **clinical treatment strategies**. Current tools require programming knowledge (e.g., MATLAB, Python).



Learn about the carbohydrate contents of different meals/food and their effect on glucose regulation.

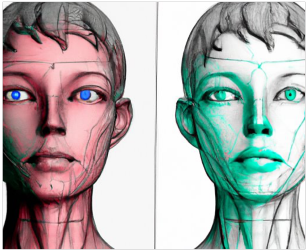
Content

Ensuring fairness and equity

Biases lead to safety issues & equitable distribution of technology.

1. Artificial pancreas systems are **high-risk medical devices**.
2. We use USA FDA-accepted **simulators/models**.
3. **Simulation / Design bias** [e,g., **populations, activity patterns** (meals/food)] → discriminate certain populations.
4. Requires **resources to build models for under-represented groups & opportunity for lived experience experts** to voice concerns.

Simulators



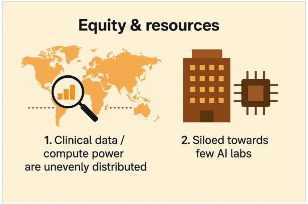


Image Credits DALL·E 3

Open-source Resources: Codebases, Tools, Datasets, and Models

CAPSML

2,050+ users, 58 countries.

(July 2023 – Oct 2025)

Visit: capsml.com

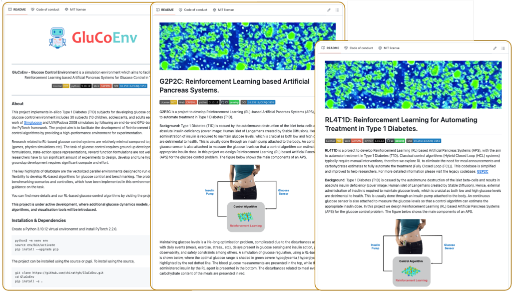
RL4H

Reinforcement Learning for Health

Repositories: G2P2C, GluCoEnv, RL4T1D

☆ 62

28

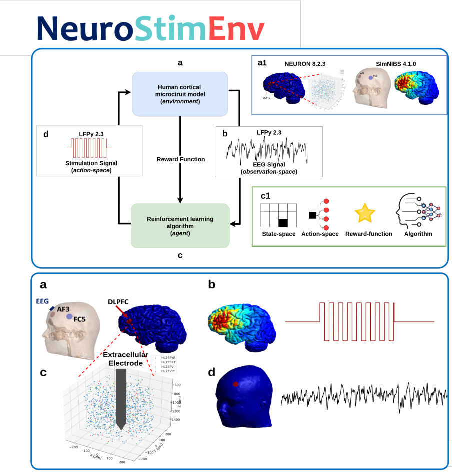


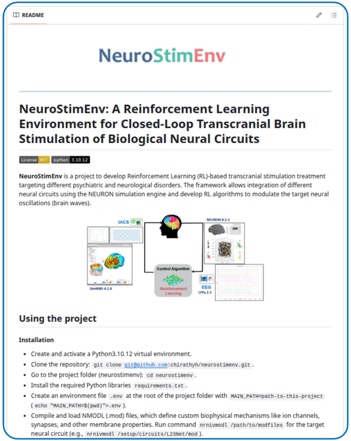
Visit: https://github.com/RL4H

Hettiarachchi, et al. "Bridging Expertise Through Open Science: A Transdisciplinary Effort to Develop Artificial Intelligence Equipped Automated Insulin Delivery Systems for Glucose Regulation in Diabetes", CHARM 2025.

Hettiarachchi, et al. "Simulated data for machine learning applications in type 1 diabetes", (in-review), 2025.

Open-source Resources: Codebases, Tools, Datasets, and Models




README

NeuroStimEnv

NeuroStimEnv: A Reinforcement Learning Environment for Closed-Loop Transcranial Brain Stimulation of Biological Neural Circuits

License MIT python 3.10.12

NeuroStimEnv is a project to develop Reinforcement Learning (RL)-based transcranial stimulation treatment targeting different psychiatric and neurological disorders. The framework allows integration of different neural circuits using the NEURON simulation engine and develop RL algorithms to modulate the target neural oscillations (brain waves).

Using the project

Installation

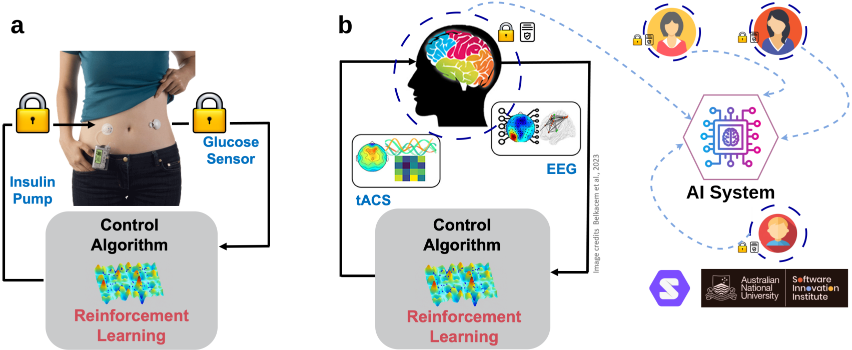
- Create and activate a Python3.10.12 virtual environment.
- Clone the repository: `git clone git@github.com:chirathyh/neurostimenv.git`.
- Go to the project folder (neurostimenv): `cd neurostimenv`.
- Install the required Python libraries `requirements.txt`.
- Create an environment file `.env` at the root of the project folder with `MAIN_PATH=path-to-this-project` (`echo "MAIN_PATH=$(pwd)">.env`).
- Compile and load NMODL (.mod) files, which define custom biophysical mechanisms like ion channels, synapses, and other membrane properties. Run command `nrnivmodl /path/to/modfiles` for the target neural circuit (e.g., `nrnivmodl /setup/circuits/L23Net/mod`).

Visit: https://github.com/chirathyh/neurostimenv

Hettiarachchi, et al. “Simulating closed-loop transcranial brain stimulation for reinforcement learning-based treatment discovery”, (in-review) 2025.

Content

Engineering solutions for consent, privacy, and security



(a) Homomorphic encryption for artificial pancreas systems and **(b)** PODs (Personal Online Data stores) provide individuals with access & personal control over their own data [work done by ANU SII].

Weng, et al., "Ensuring security of artificial pancreas device system using homomorphic encryption". Biomedical Signal Processing and Control, 79, 2023.

Content

1. Background: Reinforcement Learning (RL) & closed-loop clinical treatment
2. Shaping a future together: Clinicians, lived experience, and AI
3. Building trust through education: It's easier to trust when you understand...
4. Ensuring fairness and equity
5. Engineering solutions for consent, privacy, and security
6. **Accountability and need for regulatory frameworks**
7. Conclusion

Policy gaps and accountability

"I do understand the regulatory concern, but you have to understand that, when done correctly, and I get that's a very big question mark, ... this technology is literally life changing."
- Parent of a person with T1D [Carolyn et al., 2020, p.1].

- **Adaptive medical AI** (updates weights in deployment) is **high-risk**; can change behavior unpredictably, with extra challenges for real-time safety, reproducibility, explainability, validation, and post-market surveillance. Therefore, must meet strict risk-management and monitoring.

- **Accountability:** multi-actor (algorithm developer, cloud provider, clinic, physician, ..) **workflows/processes** blur responsibility. Focus on open standards / transparency, AI literacy/training, and designing AI-clinical workflows.

■ In T1D, patient expertise fueled by the slow pace in the development of APS, have resulted in open-source initiatives to design do-it-yourself APS. Movements, such as #WeAreNotWaiting call attention to timely solutions, value of co-creation, importance of proper regulatory frameworks [Kesavadev et al., 2020].

Kesavadev, Jothydev, et al. "The do-it-yourself artificial pancreas: a comprehensive review." Diabetes Therapy 11.6 (2020): 1217-1235.

Carolyn, J., et al. "Key Research Summary: Parents Using Unregulated Technology to Manage Type 1 Diabetes in Children". 2020. University of Melbourne, 2023.

AI in Medicine: Threats

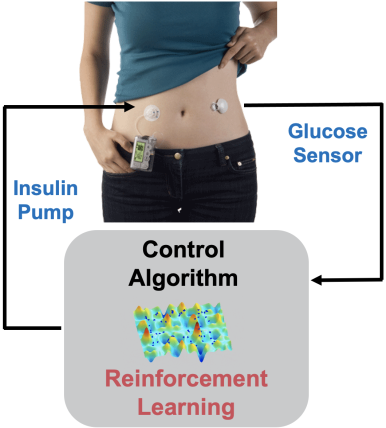
Poor practices worsens medical AI.

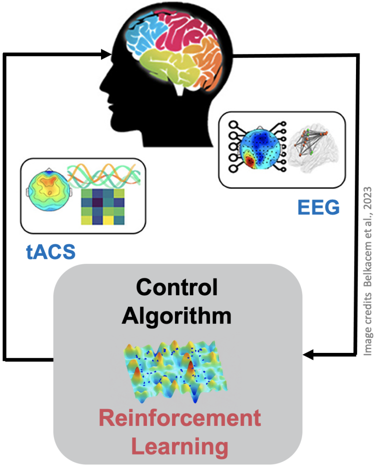
- Solutions built by **AI-experts without clinical insight**—or by **non-experts casually "vibe-coding"**—frequently produces unreliable results / sub par research.
  [Instead of using AI to do the job for you; use it to learn skills you are lacking!]
- **Irreproducible/misleading research** → **waste, lost trust / funding**.
- Poor engineering increases **security/privacy** risks.
- Repeated failures risk **regulatory backlash** or market collapse for useful AI systems.

Mitigations: Multi-disciplinary teams, external validation, prospective trials, reproducible code, continuous monitoring, education.

Content

Recap: Reinforcement Learning for Health (RL4H)





Conclusion

- **Meaningful contributions** emerge through **multi-disciplinary collaborations/research**. **The future will include AI systems integrated in the process** [Kasparov's Law].
- **Education** as a step towards improving **trust/explainability of medical AI systems**.
- Instead of using AI to do the job for you; use it to **learn the skills you are lacking!**
- **Open-science** (data, code, tools, models) for **reproducibility and benchmarking**.
- Solving these challenges take time. **Stay patient, be respectful** and keep going. . .

Thank You

Acknowledgements

- This research was funded by the MRFF 2022 National Critical Research Infrastructure (Project MRFCRI000138: "Developing a new digital therapeutic for depression: Closed loop non-invasive brain stimulation").
- This research was funded by the Australian National University (ANU), School of Computing; Our Health in Our Hands (OHIOH) strategic initiative.
- This work was supported by computational resources provided by the Australian Government through the National Computational Infrastructure (NCI) under the National Computational Merit Allocation Scheme (NCMAS), ANU Merit Allocation Scheme (ANUMAS), and ANU Startup Schemes.

Team/Collaborators

- **Diabetes (CAPSML)**: OHIOH Team; Dr David O'Neal, Dr Barbora Paldus, & Dr Dale Morrison from the Diabetes Technology Research Group, St Vincent's Hospital, Melbourne; Dr Charith Chitraranjan, University of Moratuwa (UoM), Dilshan Rakshitha; Prasanjith Lorensuhewa, Damika Anupama (Honours, UoM); David Timms (Masters, ANU), Jordan Triming (Honours, ANU), Samuel Price (Honours, ANU), Sam Cantrill, Dr Nicolo Malagutti, Prof. Christoper Nolan, Dr Elena Daskalaki, Dr Chirath Hettiarachchi, Prof. Hanna Suominen.
- **Depression-tACS**: Project Team [Hardware, Clinical, ANU SII]; Dr Chirath Hettiarachchi, Dr Neil Bailey, Prof. Paul B Fitzgerald, Prof. Hanna Suominen.