What *we* need to know about AI and healthcare

Vlada Rozova, PhD
Senior Research Fellow
Centre for Digital Transformation of Health
University of Melbourne

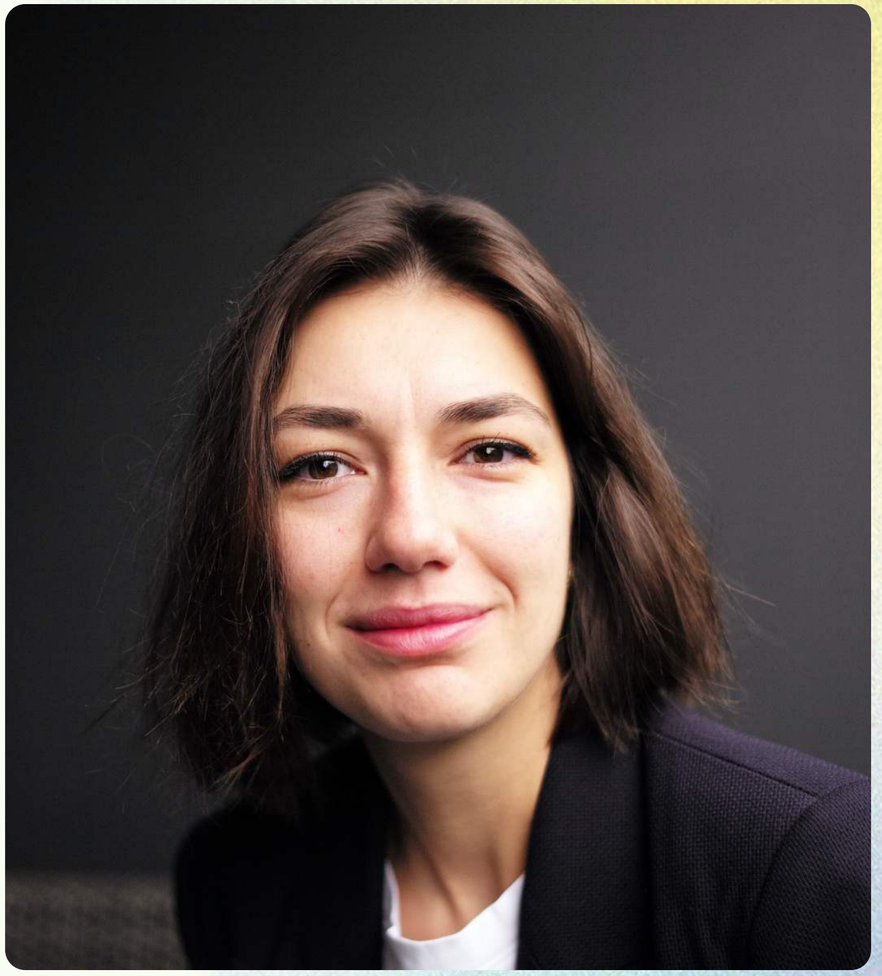
THE UNIVERSITY OF MELBOURNE

About me

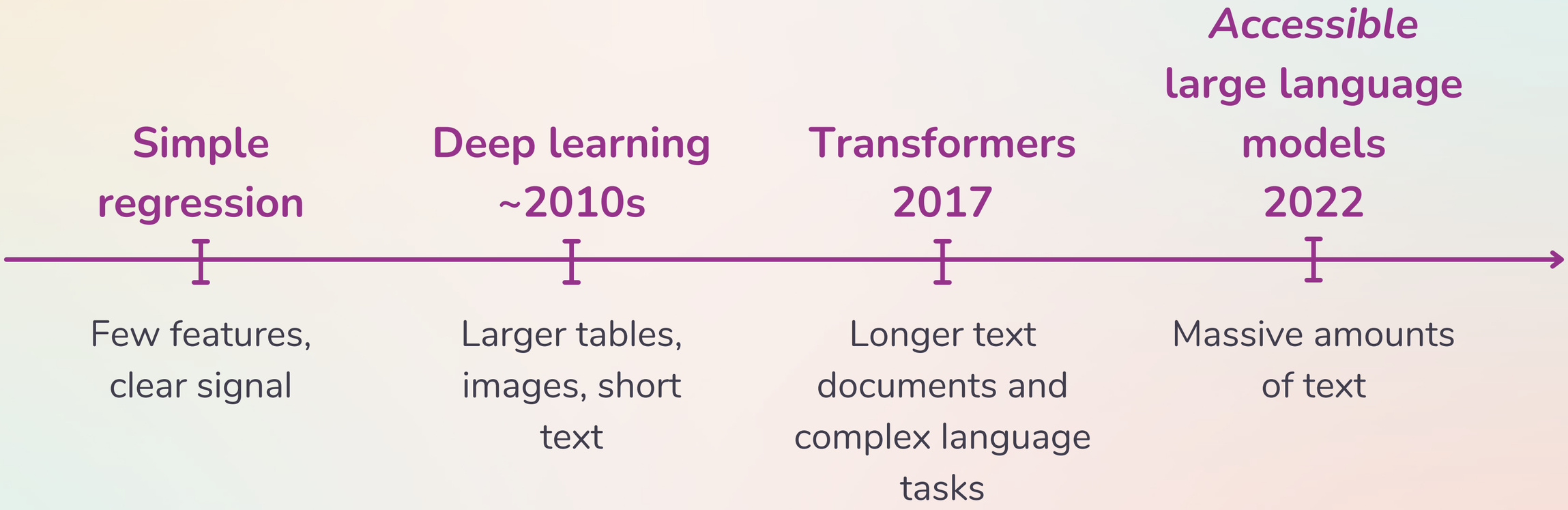
- Senior Research Fellow in Applied Machine Learning
- BSc and MSc in Applied Mathematics & Computer Science
- PhD in Machine Learning in Biomedicine: Cancer cell behaviour in response to external factors
- Research interests & expertise: clinical AI, data infrastructure, AI delivery in health

Email: vlada.rozova@unimelb.edu.au

LinkedIn

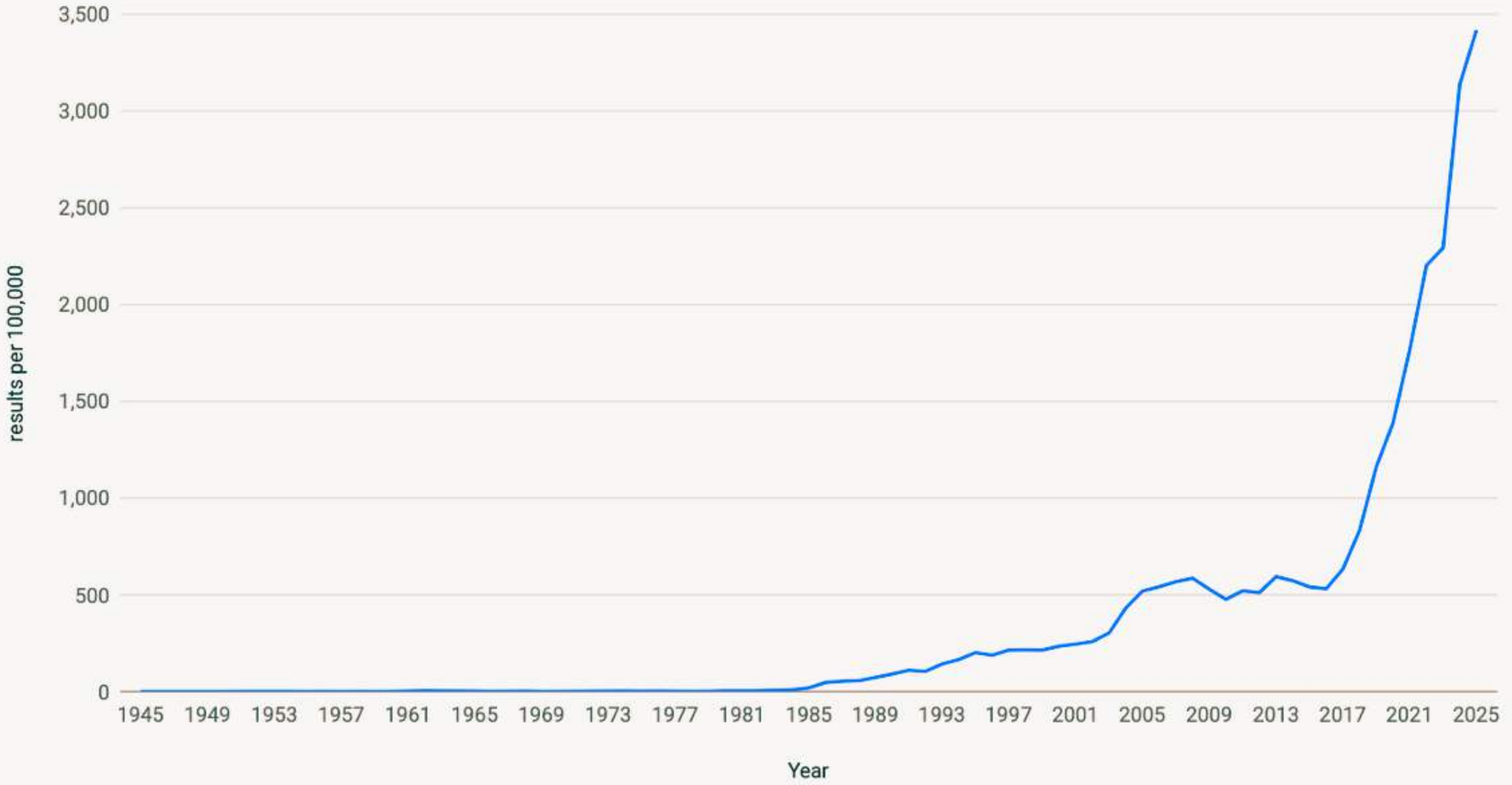


How did we get here?



Results per 100,000 citations in PubMed

proportion for each search by year, 1945 to 2025



The next cool thing...

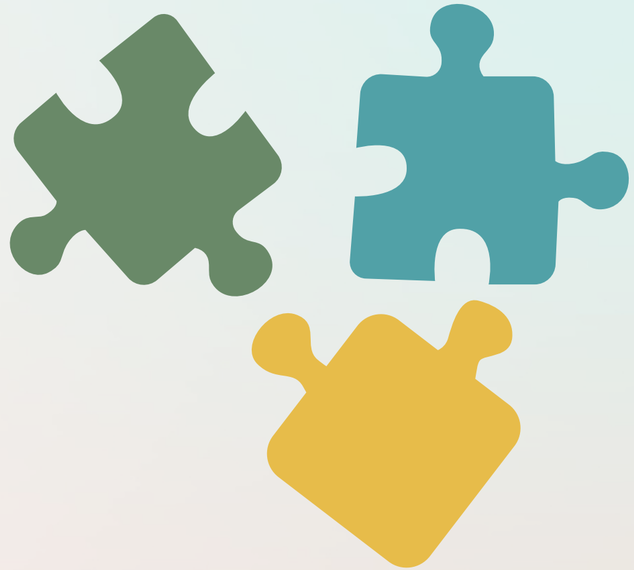
Supervised learning / discriminative AI

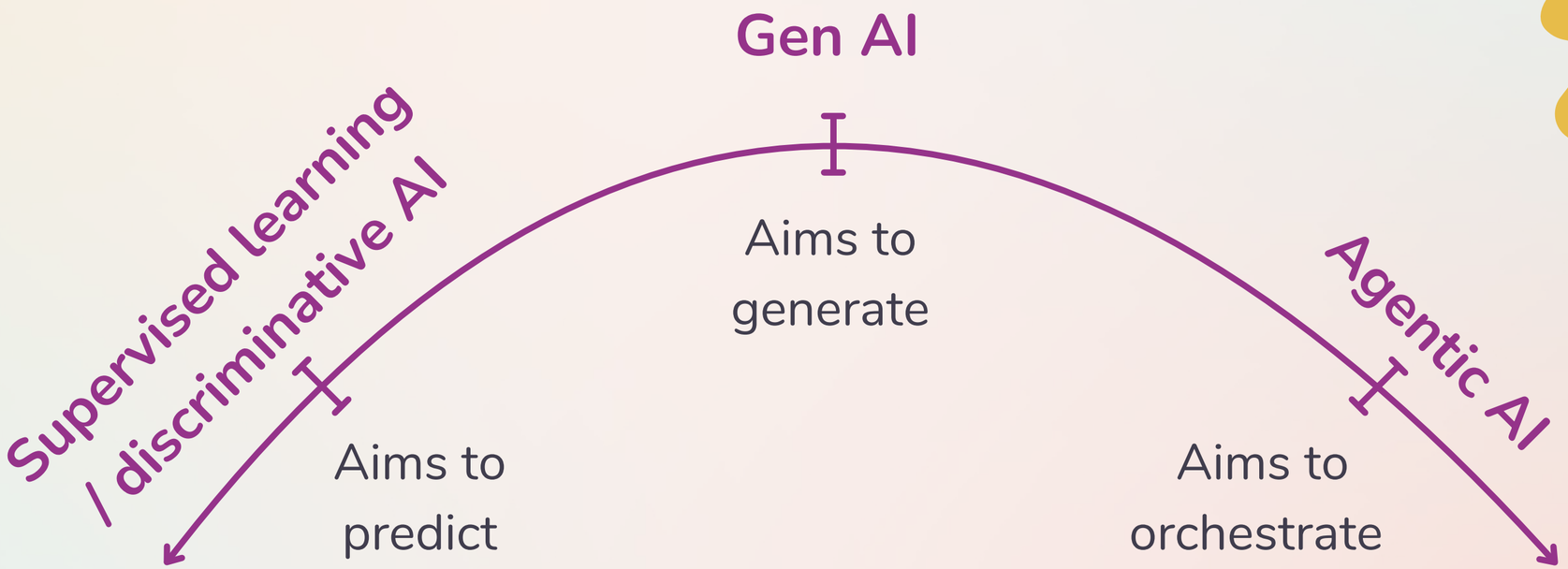
Gen AI

Agentic AI



Not a linear progression





AI as diagnostic intervention



AI as clinical assistant



AI as diagnostic intervention



Overwhelmingly supervised learning

Quality of inputs

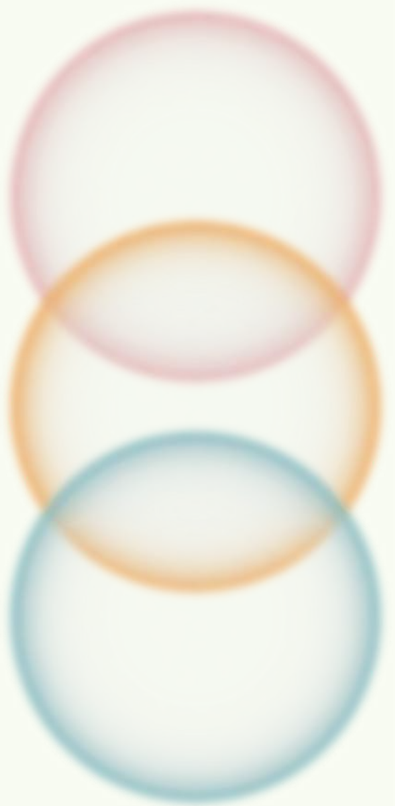
Quality of outputs



Well-formulated problem



Valuable outputs



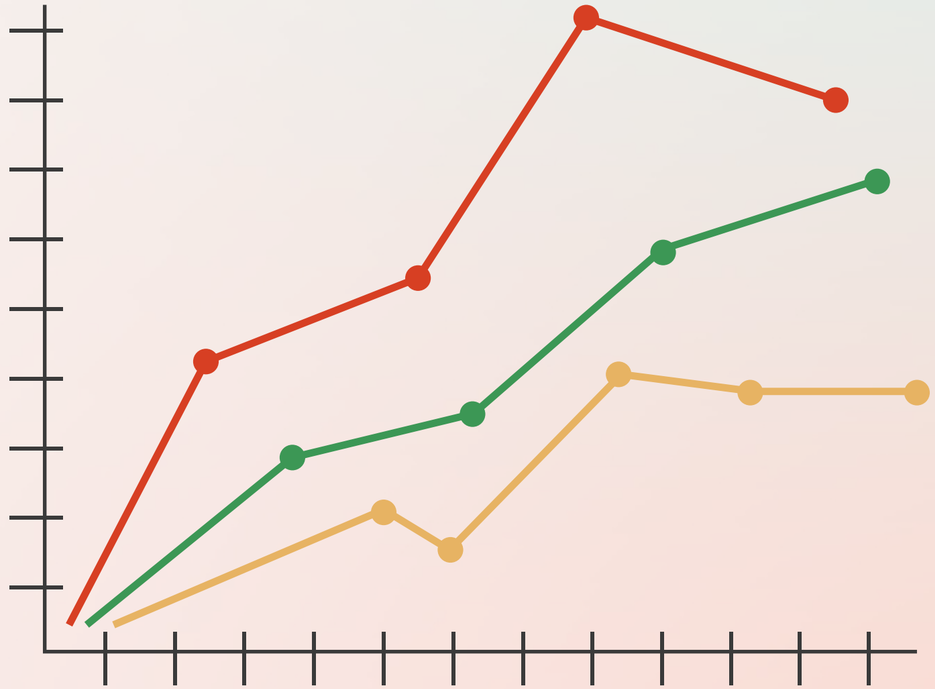
Application context

Clinical vs epidemiological

- Risk prediction – *“risk that x will happen”*
- Classification/detection – *“patient has condition y”*

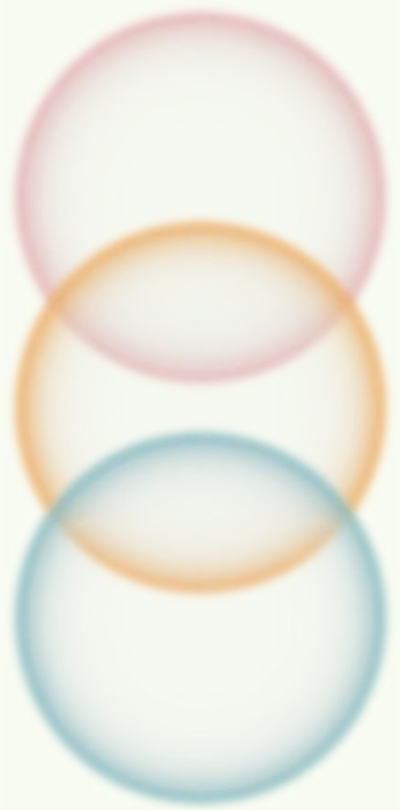
Actual	Predicted: Negative	Predicted: Positive
Negative	True negatives: 80,000	False positives: 200
Positive	False negatives: 200	True positives: 500

- May not be sufficient for individual patients
- Good enough to grasp the overall trend



Use case determines the primary metrics



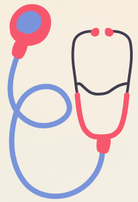


Model accuracy

≠

Clinical utility

How inadequate labels may lead to a clinically meaningless model



Admission

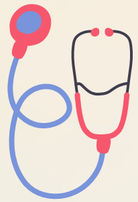


Discharge

30-day readmission?

Why hospital readmissions matter:

- Patients do not want to come back to the hospital
- Hospitals are paid less
- Readmissions are expensive for payers



Admission



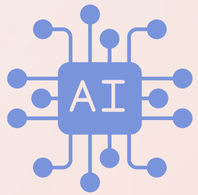
Discharge

30-day readmission?

24h before discharge



Patient data up to 24h before discharge



AI model trained to predict readmissions



Admission



AI prediction



Discharge

30-day readmission?

24h before discharge

Opportunity for intervention

For example:

- Organise patient follow-up
- Set up patient monitoring
- Provide patient education



AI prediction

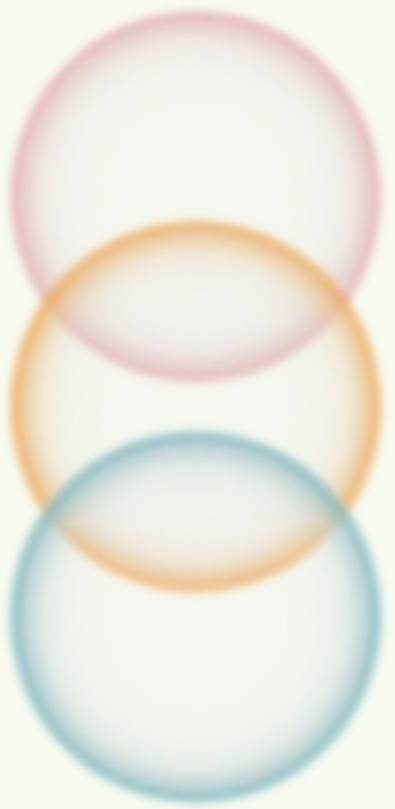
Clinician anticipates readmission

Preventative actions taken	Yes	No
Yes	0%	0%
No	✓	✓

← **Only beneficial** where there was no intervention

Clinical workflow
is not
an afterthought





One-off validation

≠

License for continuous use

"There is no such thing as a validated model"

Van Calster, et al. There is no such thing as a validated prediction model. BMC Med 21, 70 (2023).

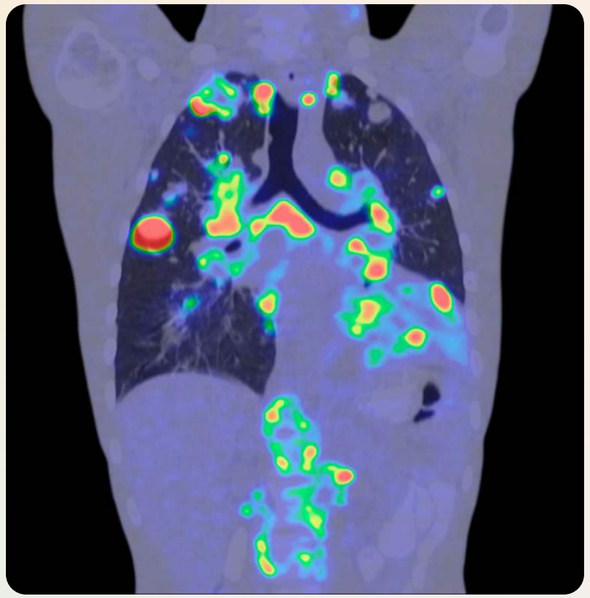
High-quality local annotated data *is*



Because “gold” data:

- Is the only objective way to verify that a commercial product performs as advertised
- Allows benchmarking across settings and systems
- Helps detect data drift
- Helps identify biases to ensure equitable care

PET scan



Report

“Clinical notes: ...Persistent fevers and rigors with no localising source or symptoms.

...

Mild diffuse peri-dental uptake...Minimal uptake in a small ground glass opacity in the right lung apex...unlikely to explain the clinical presentation...There is mild focal uptake in a cutaneous lesion.

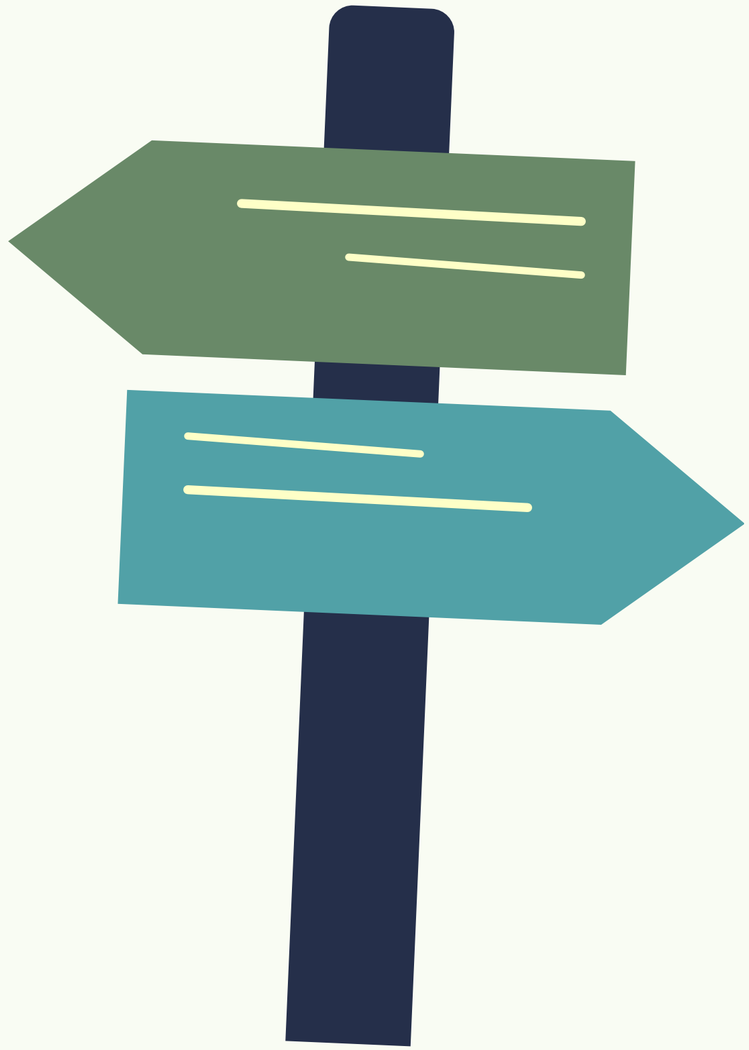
...

Conclusion: There is no PET/CT evidence of active infection....

↓

- **Radiologist:** report clearly states there is *no* active infection
- **Infectious diseases physician:** infection *cannot be fully ruled out* and is possible given the clinical presentation

Data annotation is clinical decision-making



AI as clinical assistant

Reframing for autonomy



Levels of autonomy

Human-to-Machine (H2M) Scale™



Level 0

All Human

All human, all the time.
The system does not use any AI and is only capable of doing what it has been instructed to do All automation is manual.



Level 1

Mostly Human

Limited intelligent automation.
The system uses AI in certain aspects, but it is largely reliant on human inputs and oversight.



Level 2

Half & Half

Half human, half machine.
The system can manage most aspects of the use case, but still requires human inputs and oversight.



Level 3

Mostly Machine

Predominantly AI powered.
The system can operate without human inputs or oversight in select conditions.



Level 4

All Machine

Full autonomy.
The system can perform at or above human level without inputs or oversight. The human simply defines the desired outcome, and the machine does all the work.

https://www.marketingaiinstitute.com/blog/introducing-the-human-to-machine-scale-for-rating-marketing-ai-technology

Levels of autonomy

Human-to-Machine (H2M) Scale™



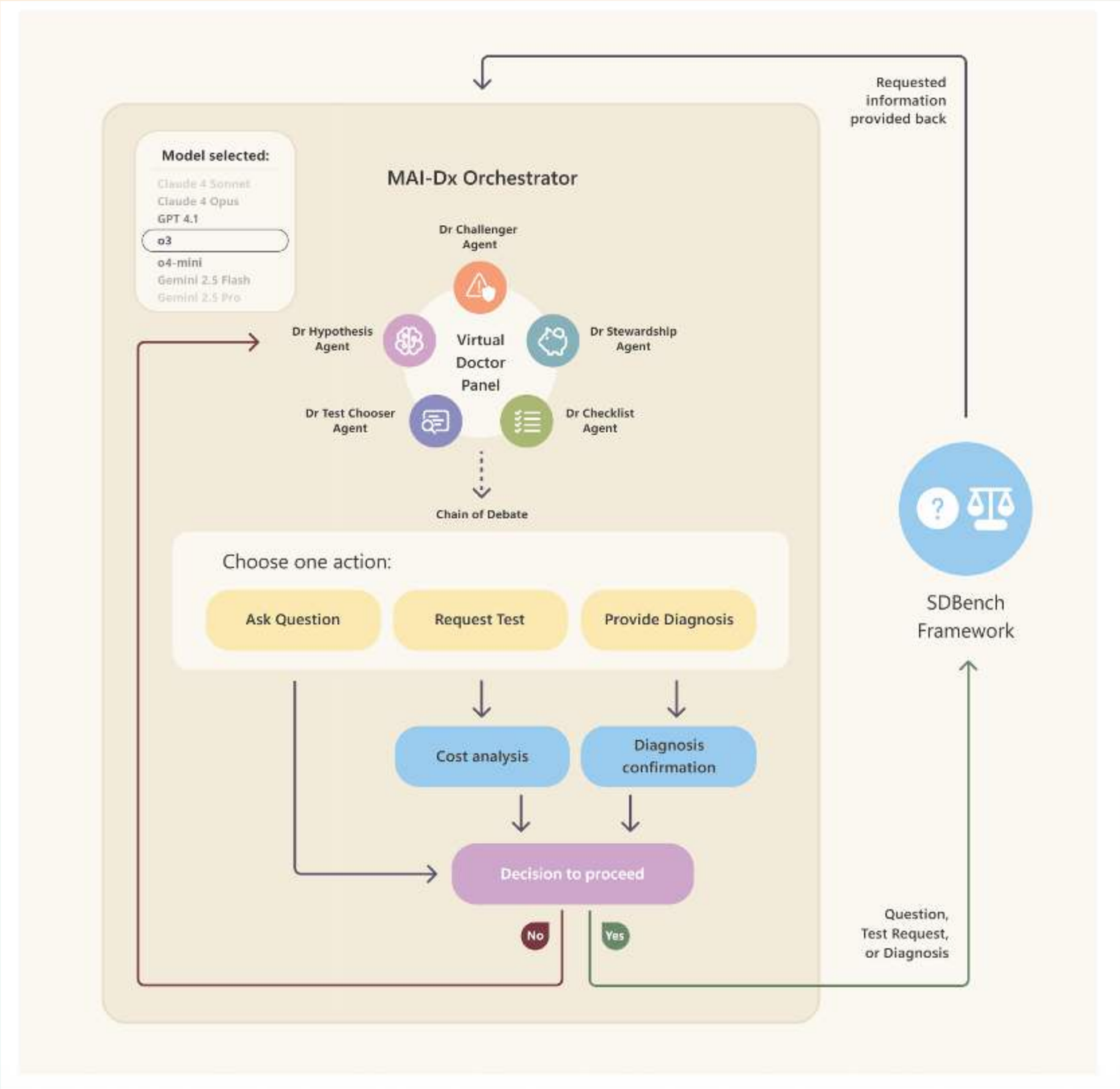
Level 0	Level 1	Level 2	Level 3	Level 4
All Human	Mostly Human	Half & Half	Mostly Machine	All Machine
No assistance	Create the subject line	Personalise based on user interests	Plan the content	Monitor performance

https://www.marketingaiinstitute.com/blog/introducing-the-human-to-machine-scale-for-rating-marketing-ai-technology

In healthcare

- Searching, synthesising, summarising, creating new content
- Retrieving and applying guidelines
- Comparing, extracting, verifying, analysing data
- For complex tasks: agentic AI framework

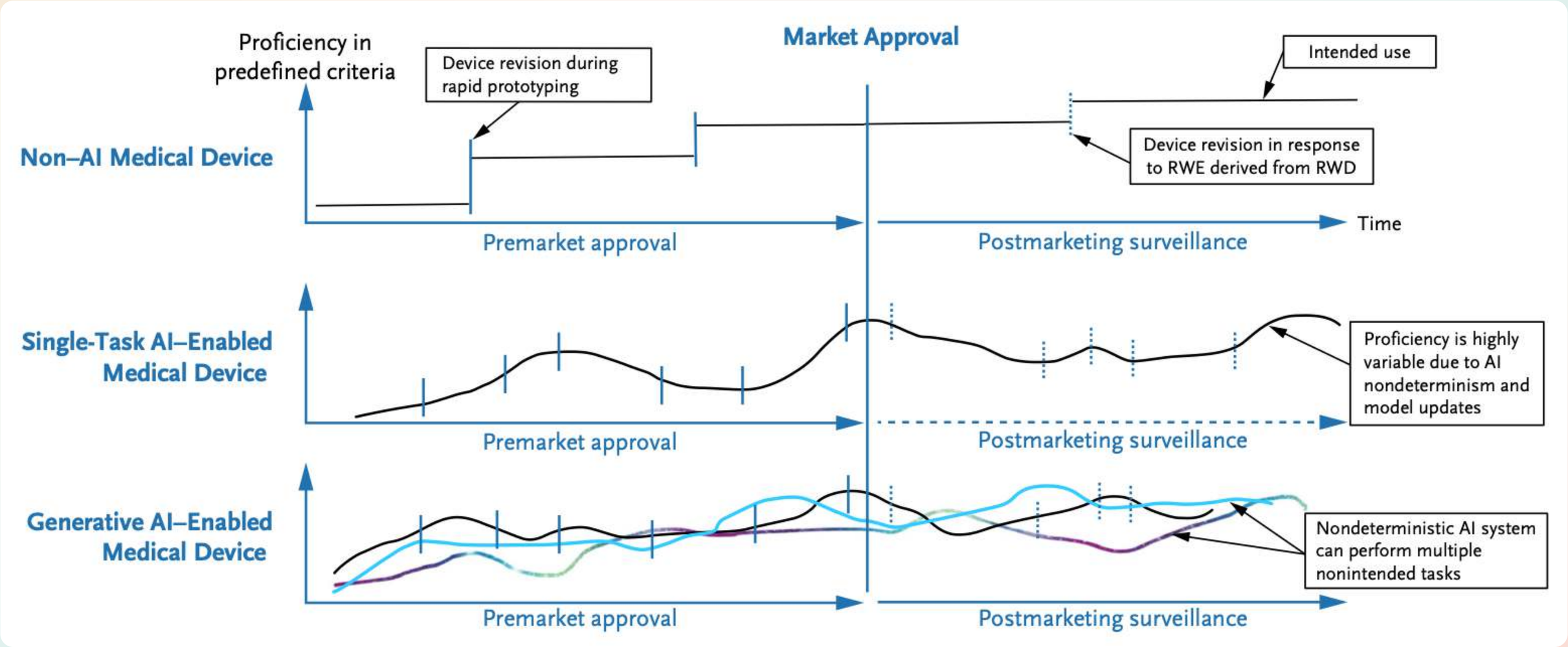
Spoiler: this is separate from managing emails and calendars with Microsoft Copilot



- Simulates a panel of physicians, proposes likely differential diagnoses, and selects high-value, cost-effective tests
- Evaluated on 304 diagnostically challenging NEJM cases
- A framework for breaking down complex reasoning tasks into competency domains

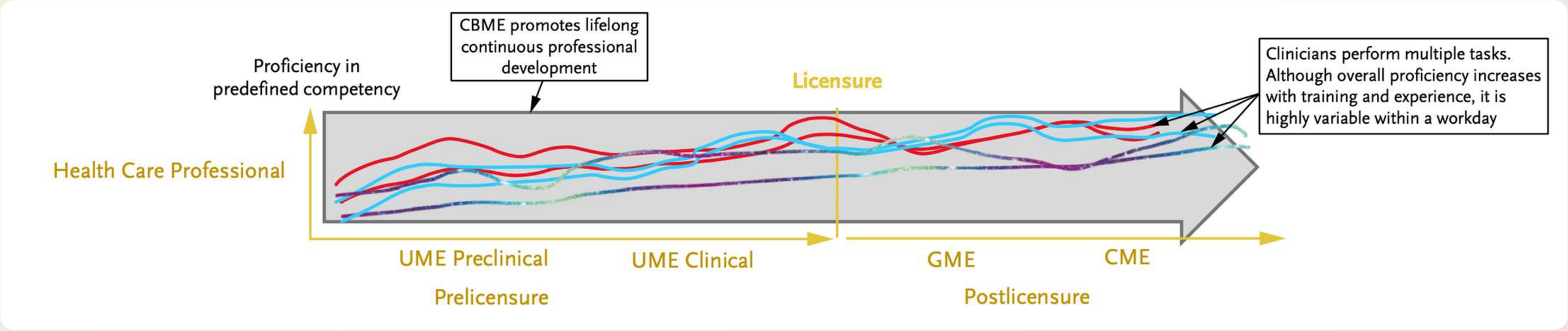
Nori, Harsha, et al. "Sequential diagnosis with language models." arXiv preprint arXiv:2506.22405 (2025).

Improvement is not guaranteed



Vokinger KN, Soled DR, Abdulnour RE. Regulation of AI: Learnings from Medical Education. NEJM AI. 2025 May;2(5).

Lifelong professional development



↑

We wouldn't give *any* task to *any* clinician...

Why do we expect a pile of maths to be good at *everything*?

Vokinger KN, Soled DR, Abdulnour RE. Regulation of AI: Learnings from Medical Education. NEJM AI. 2025 May;2(5).

From single-point, accuracy-based assessment



To an ongoing cycle of improvement & re-evaluation based on specific clinical competencies

Thank you!

Email: vlada.rozova@unimelb.edu.au

LinkedIn

Collaborating to realise the power of **data, AI, and digital technology** in healthcare

Centre for Digital Transformation of Health

THE UNIVERSITY OF MELBOURNE

Visit us:
Level 8, Melbourne Connect
700 Swanston Street, Carlton
Victoria, 3053, Australia

Sign up for updates:



MACH
Melbourne Academic
Centre for Health

2026 Research Design Webinar Series



Survey for feedback