SAFETY ASSESSMENT OF THE AI CLINICIAN FOR SEPSIS TREATMENT: AN ITERATIVE APPROACH

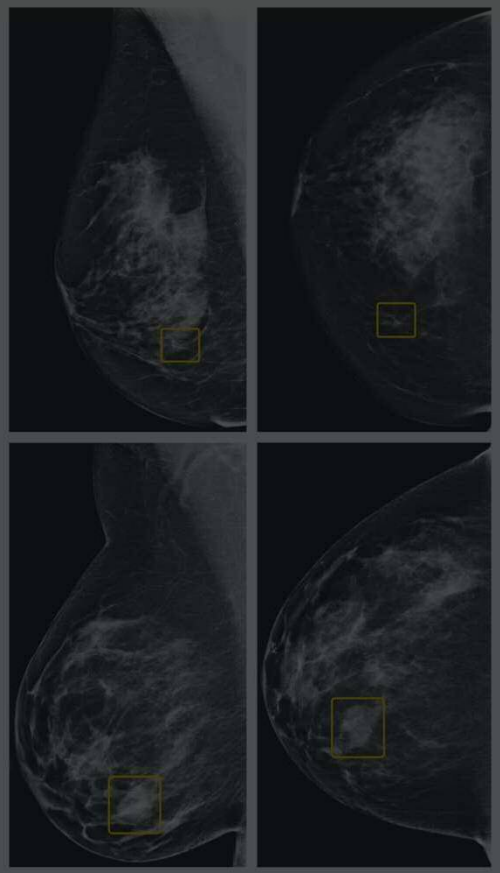
05/07/2022

Paul Festor - Future of Health 2022
Project subervised by A. Aldo Faisal & Matthieu Komorowski

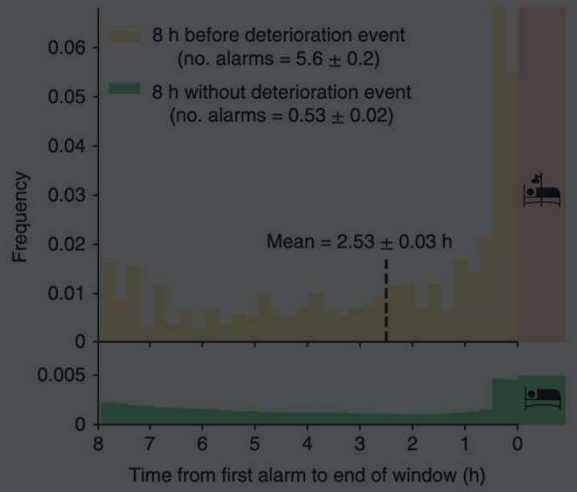
You said CDSS ? (Clinical Decision Support System)

AI for Clinical Decision Support

What ? Diagnosis



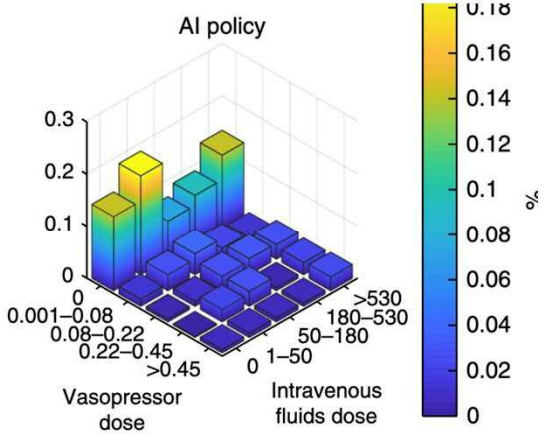
McKinney, Scott Mayer, et al. "International evaluation of an AI system for breast cancer screening." *Nature* 577.7788 (2020): 89-94.



When ? Organ failure prediction

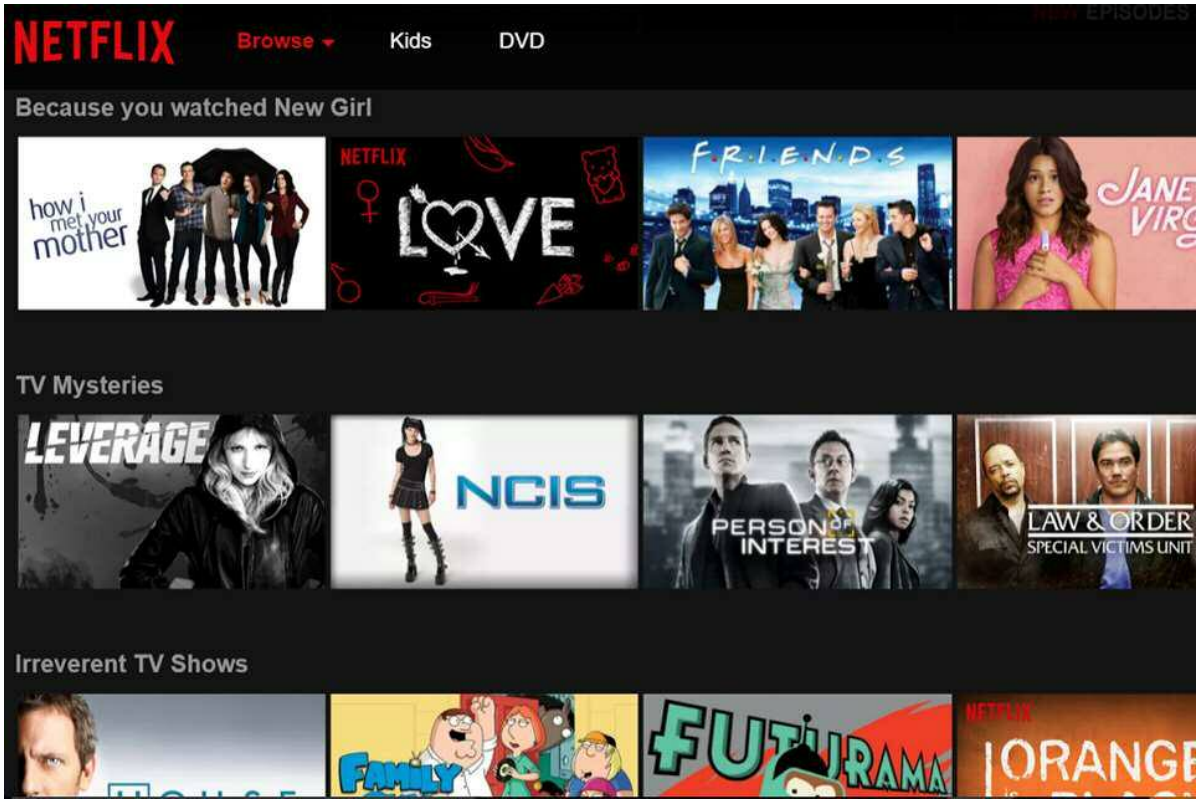
Hyland, Stephanie L., et al. "Early prediction of circulatory failure in the intensive care unit using machine learning." *Nature medicine* 26.3 (2020): 364-373.

How ? Treatment recommendation



Komorowski, Matthieu, et al. "The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care." *Nature medicine* 24.11 (2018): 1716-1720.

Recommendation systems ?



Usage data

Recommentation engine (AI)

Recommended shows

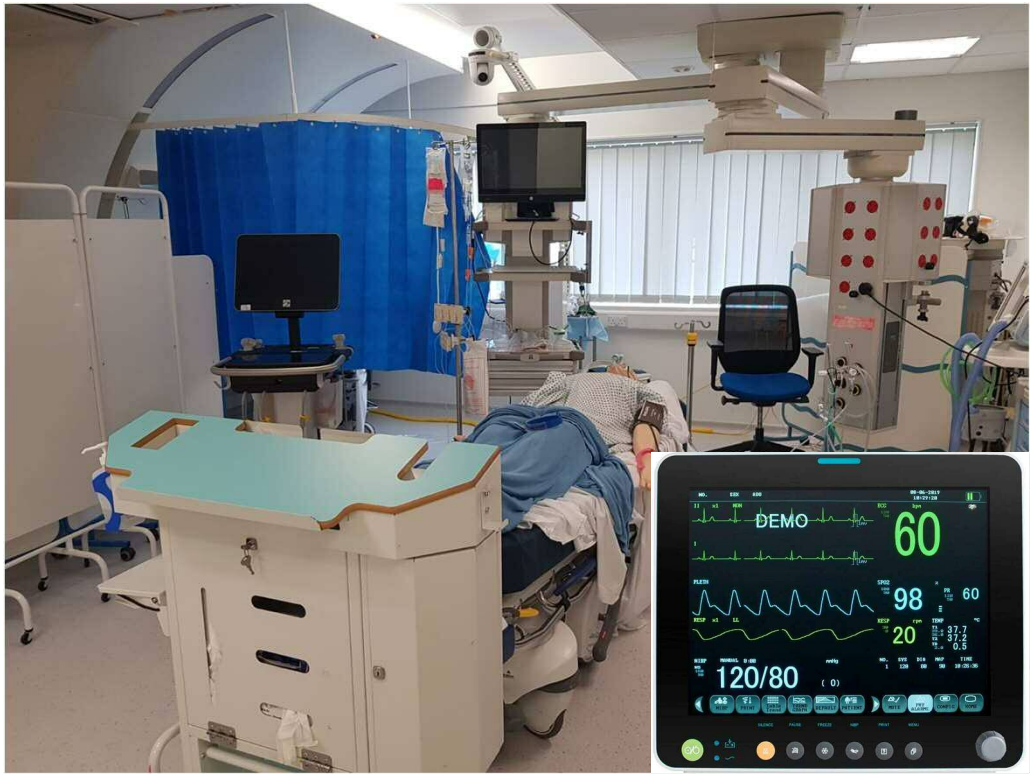
Recommendation systems ?

Patient data

Recommendation engine (AI)

Recommended treatment

Intensive Care Units: massive sources of healthcare data time series



MIMIC

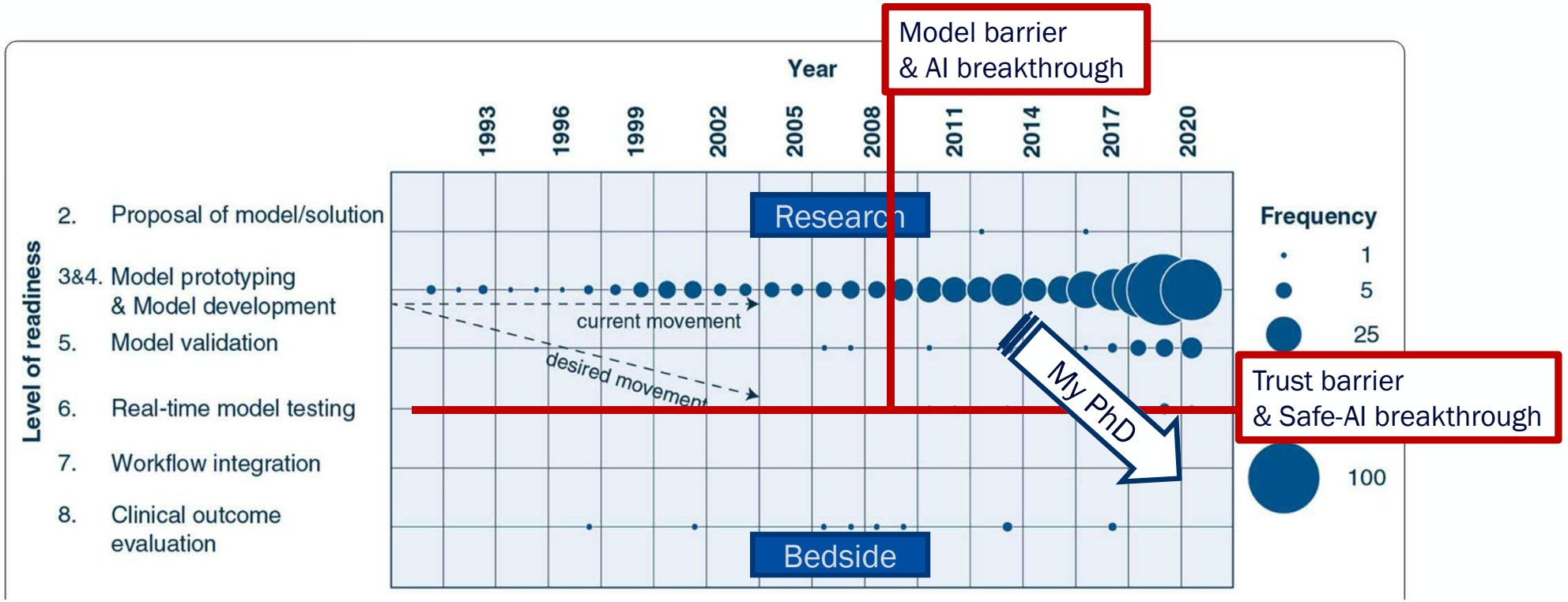
Amsterdam UMC
Universitair Medische Centra

eICU Collaborative Research Database

NHS
Imperial College Healthcare
NHS Trust

- Hundreds of ICUs
- Millions of recorded patient stays

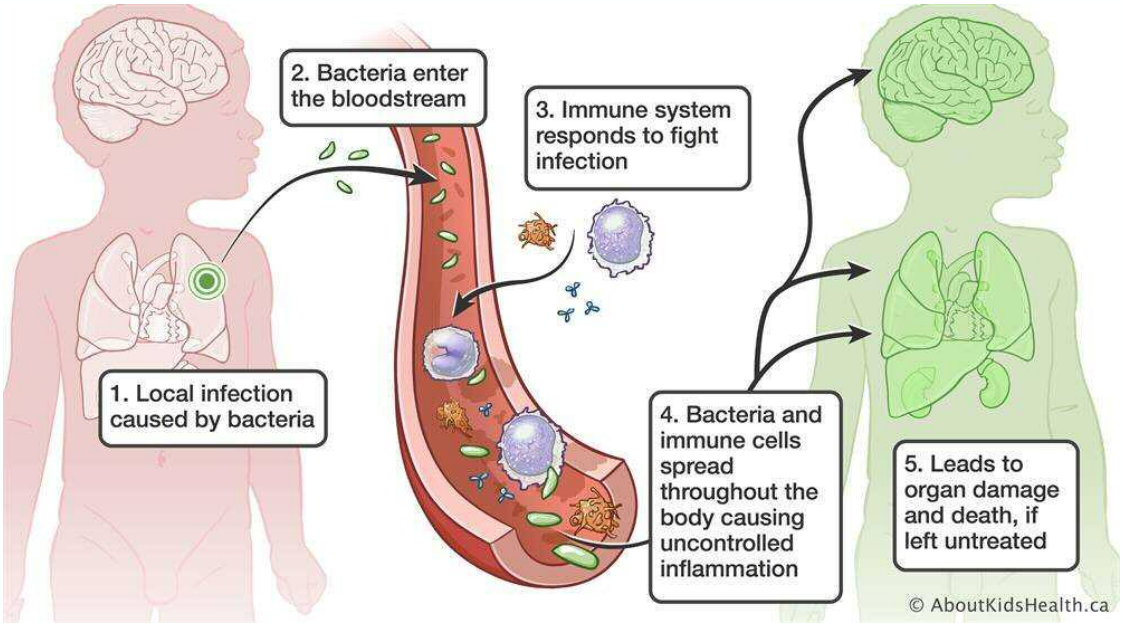
CDSSs: research/bedside gap



→ Major barrier: concerns for patient safety

Van de Sande, Davy, et al. *Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit*, 2021

Sepsis: a life-threatening disease



2017	49 million people affected worldwide
	11 million deaths worldwide
	£15.6 billion / year cost for the NHS

- Treatment: antibiotics + help organ functions long enough

WHO, *Global report on the epidemiology and burdon of* sepsis, 2020, YHEC, *The cost of sepsis care in the UK*, 2017

Treatment: multiple organ support

Mechanical ventilation

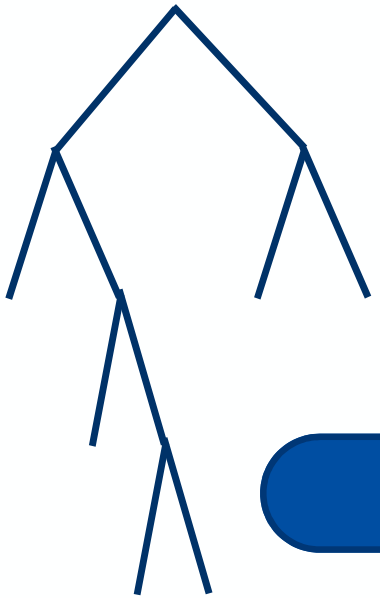
Dialysis machine

IV fluids and vasopressors

AI Clinician focus

The AI Clinician: Reinforcement Learning

- Evolves in an **environment**, learns **actions** to optimise a **reward**

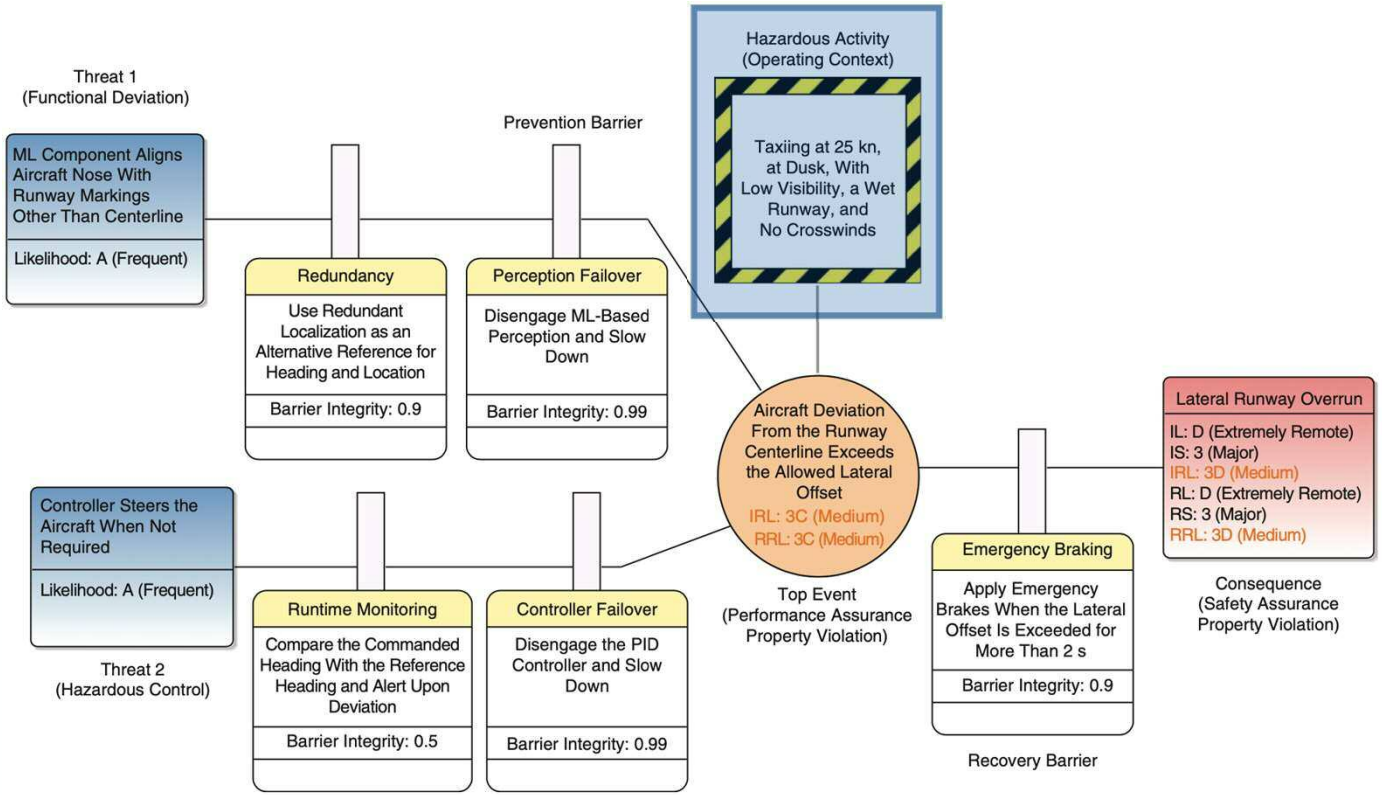


→ How to assess the safety of such systems ?

Komorowski, Matthieu, et al. "The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care." *Nature medicine* 24.11 (2018): 1716-1720.

Adding safety evidence in an assurance case: example of a taxying system

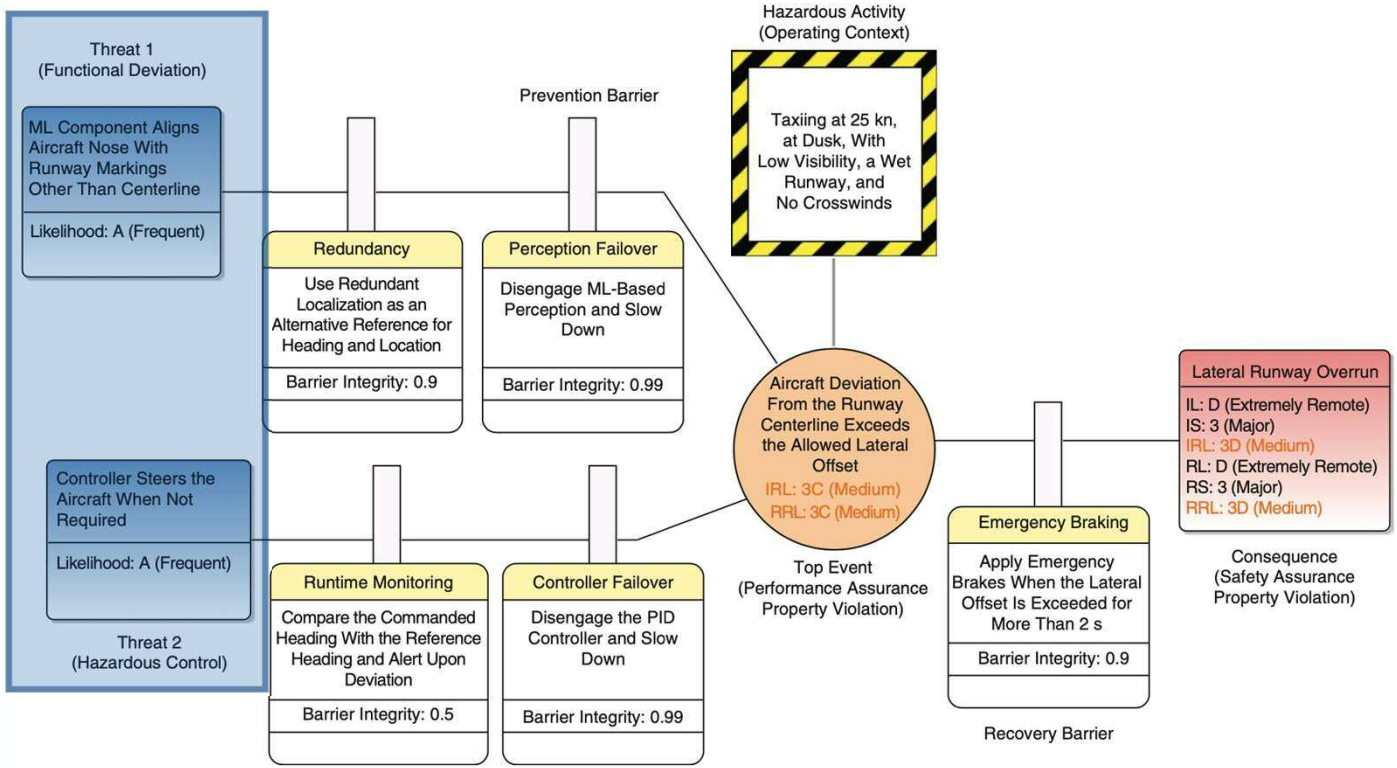
- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Adding safety evidence in an assurance case: example of a taxying system

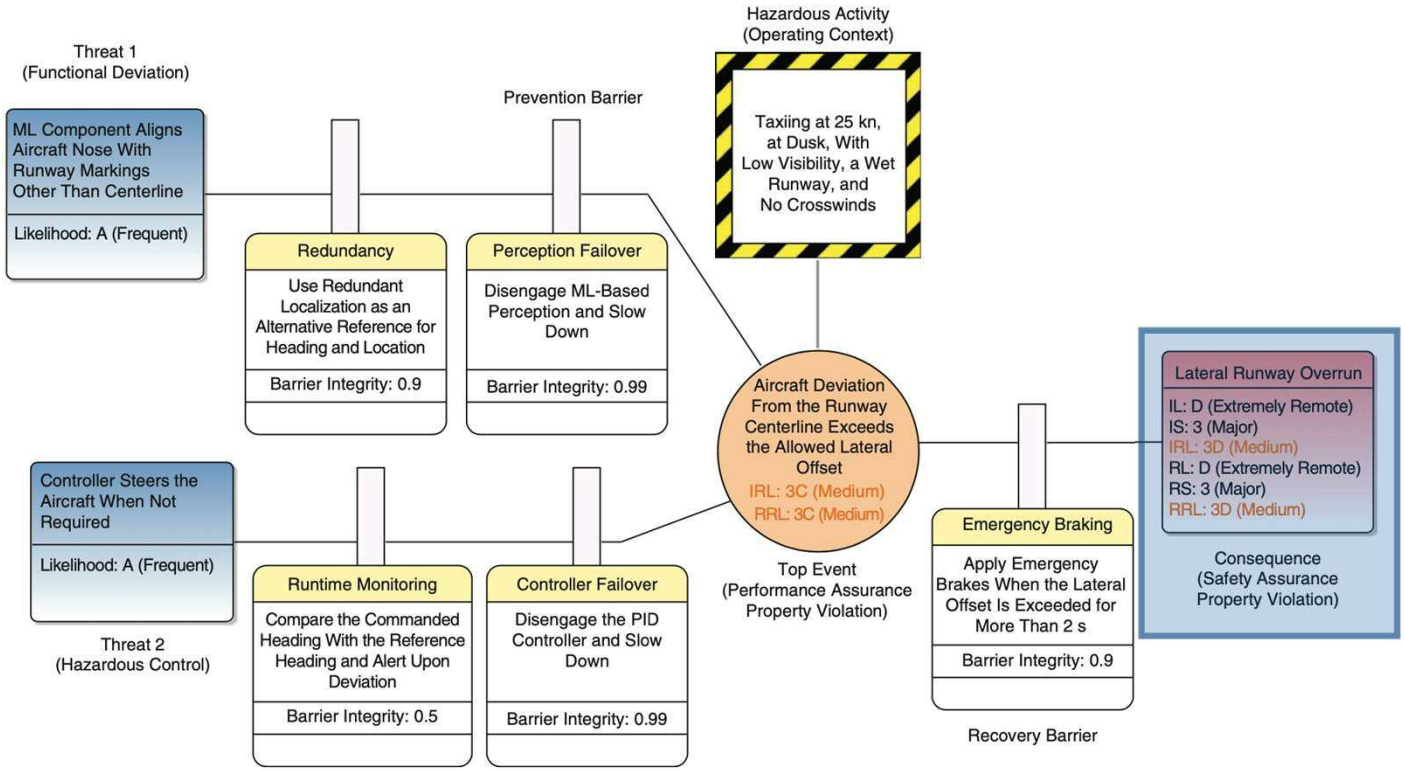
- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Adding safety evidence in an assurance case: example of a taxying system

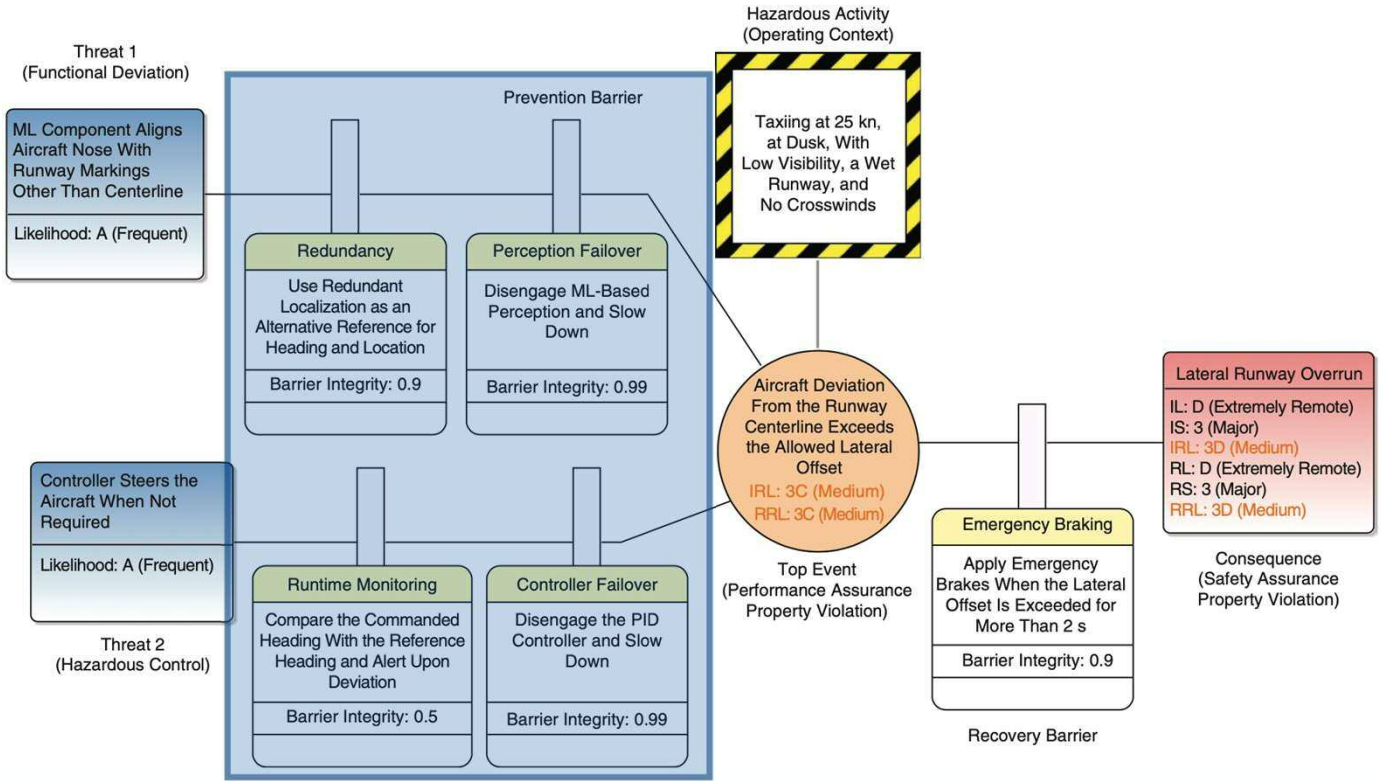
- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Adding safety evidence in an assurance case: example of a taxying system

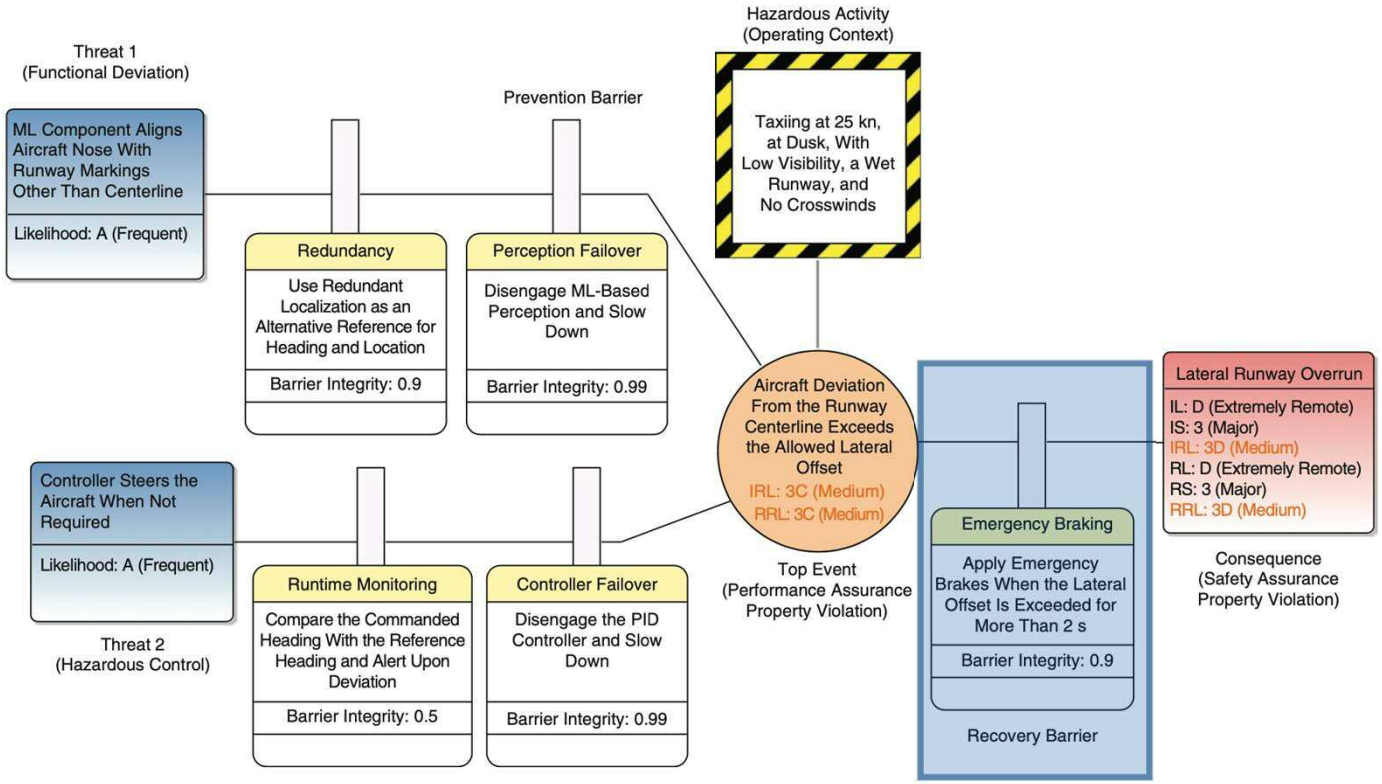
- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Adding safety evidence in an assurance case: example of a taxying system

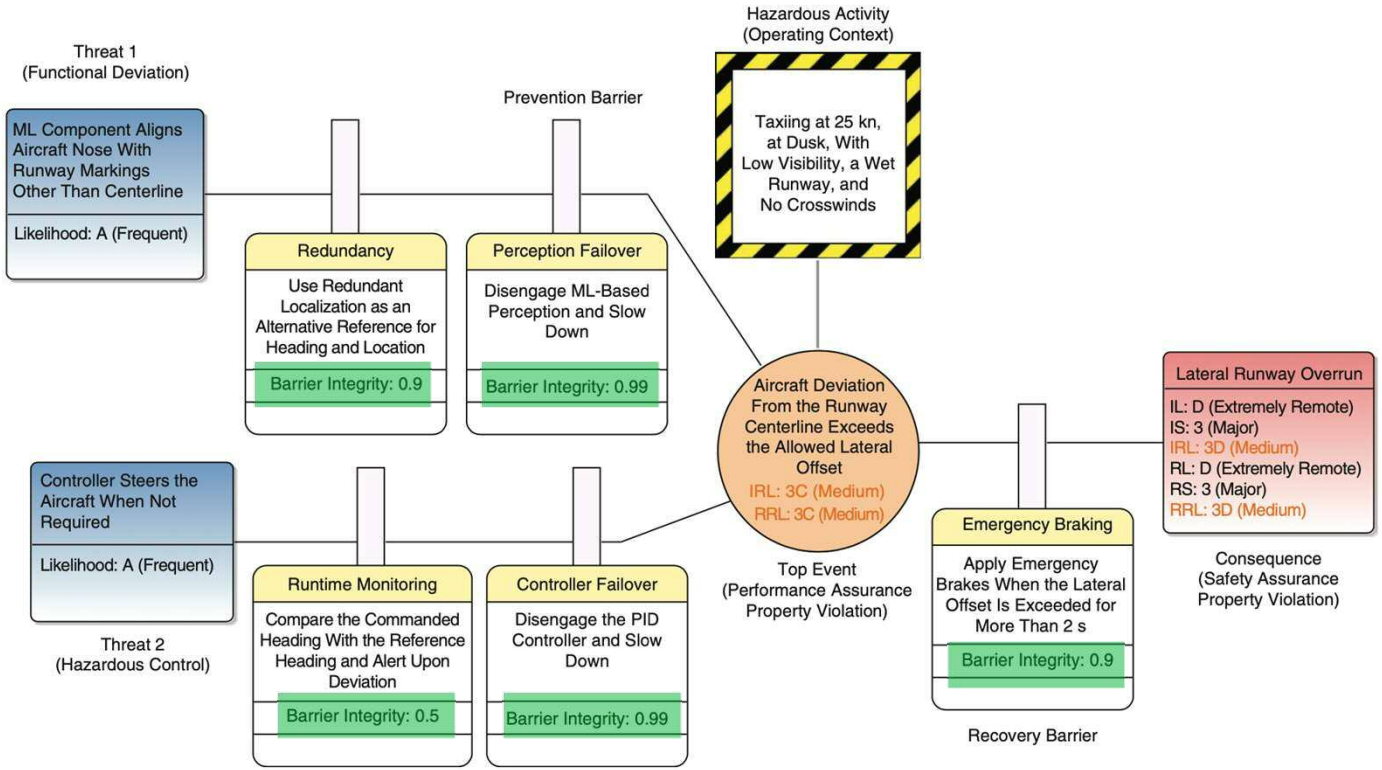
- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

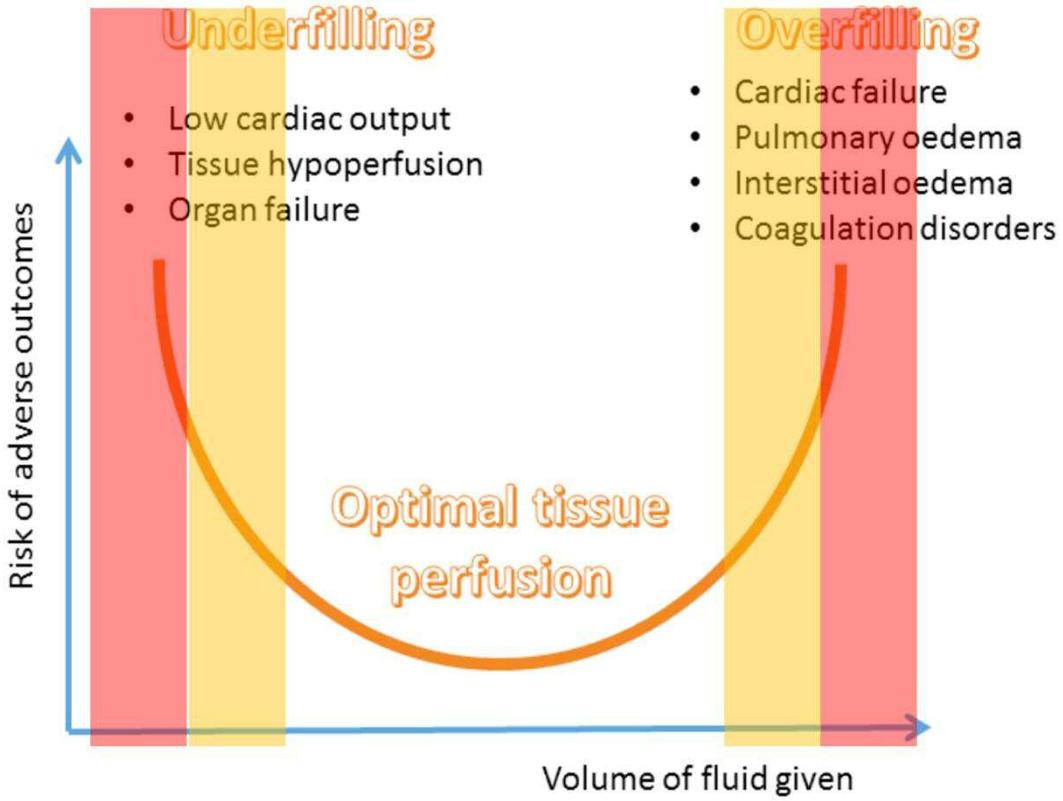
Adding safety evidence in an assurance case: example of a taxying system

- **Safety assurance case:** The document that demonstrates the overallsafety level of a system



Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Rule-based safety assessment: what is “unsafe” ?

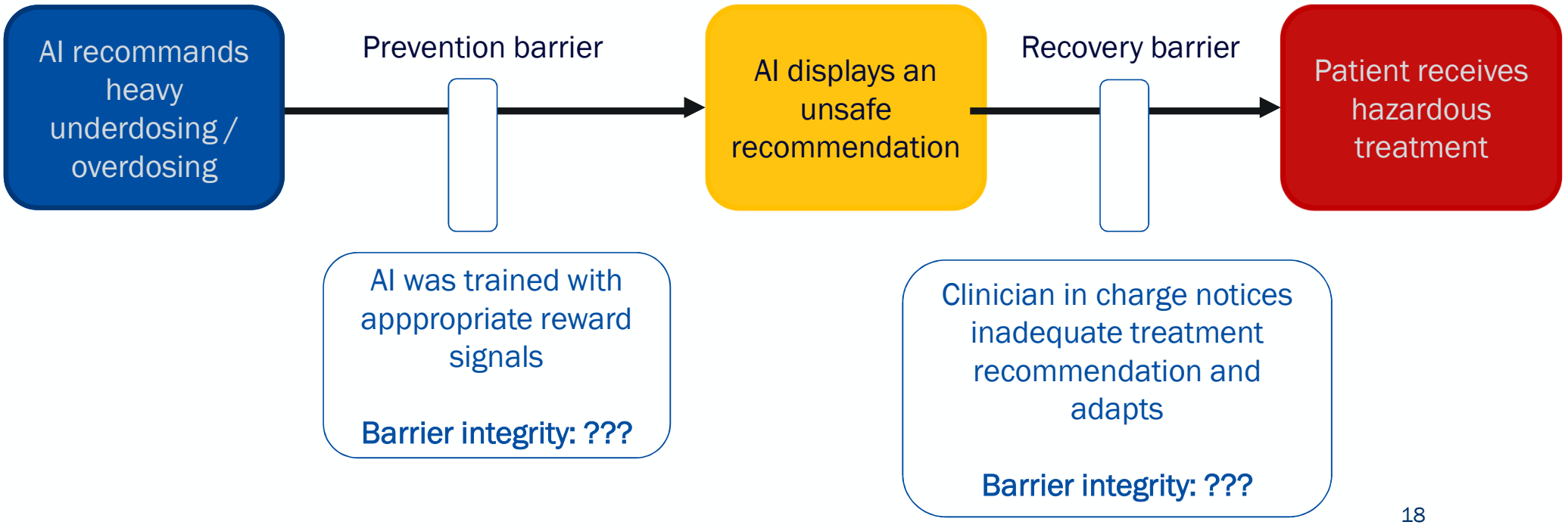


- No Vasopressors when BP is low
- High Vasopressors when BP is high

- No fluids when dehydrarted
- High fluids when fluid overloaded

Underdosing/overdosing hazard analysis diagram

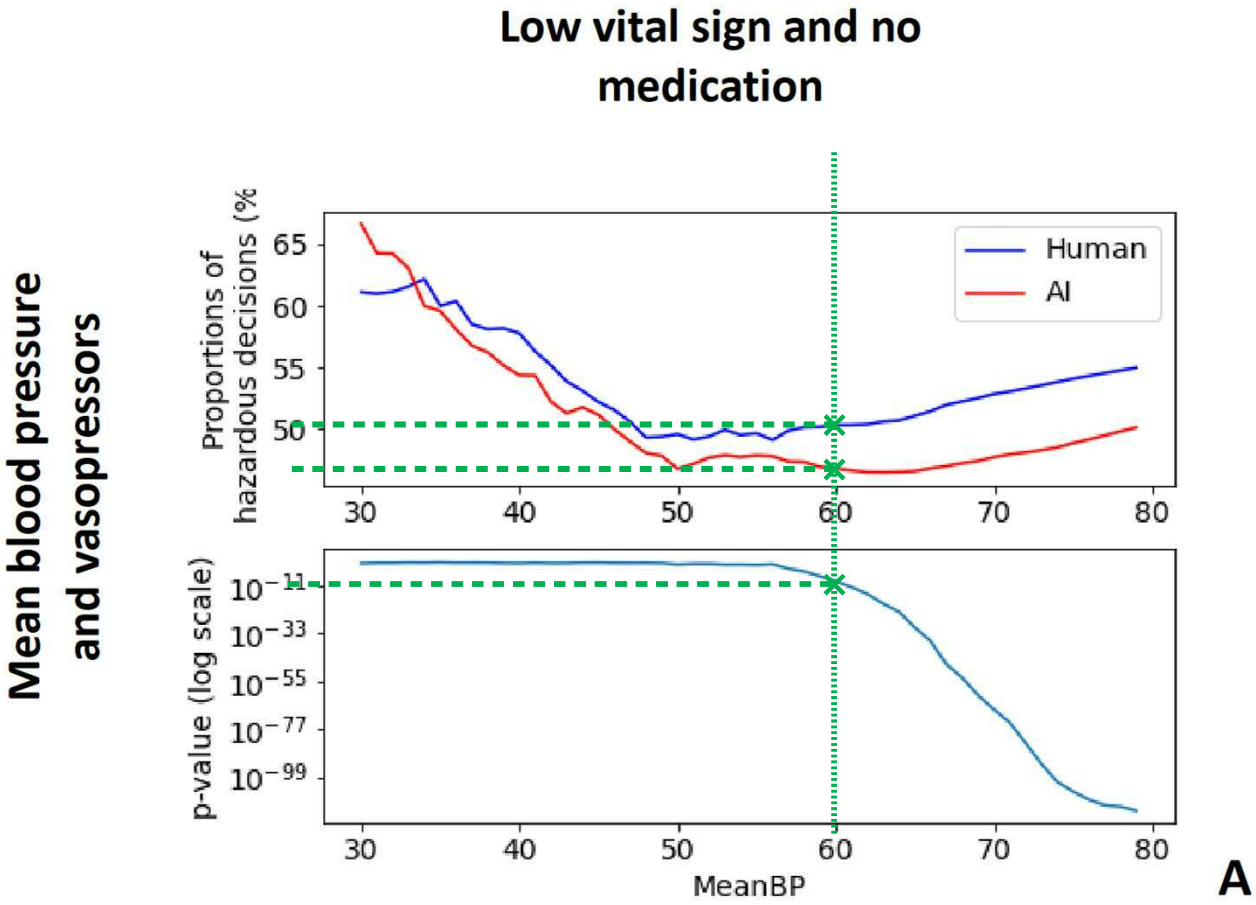
- Given a patient in one of the 4 hazardous scenarios (16% of our MIMIC cohort)



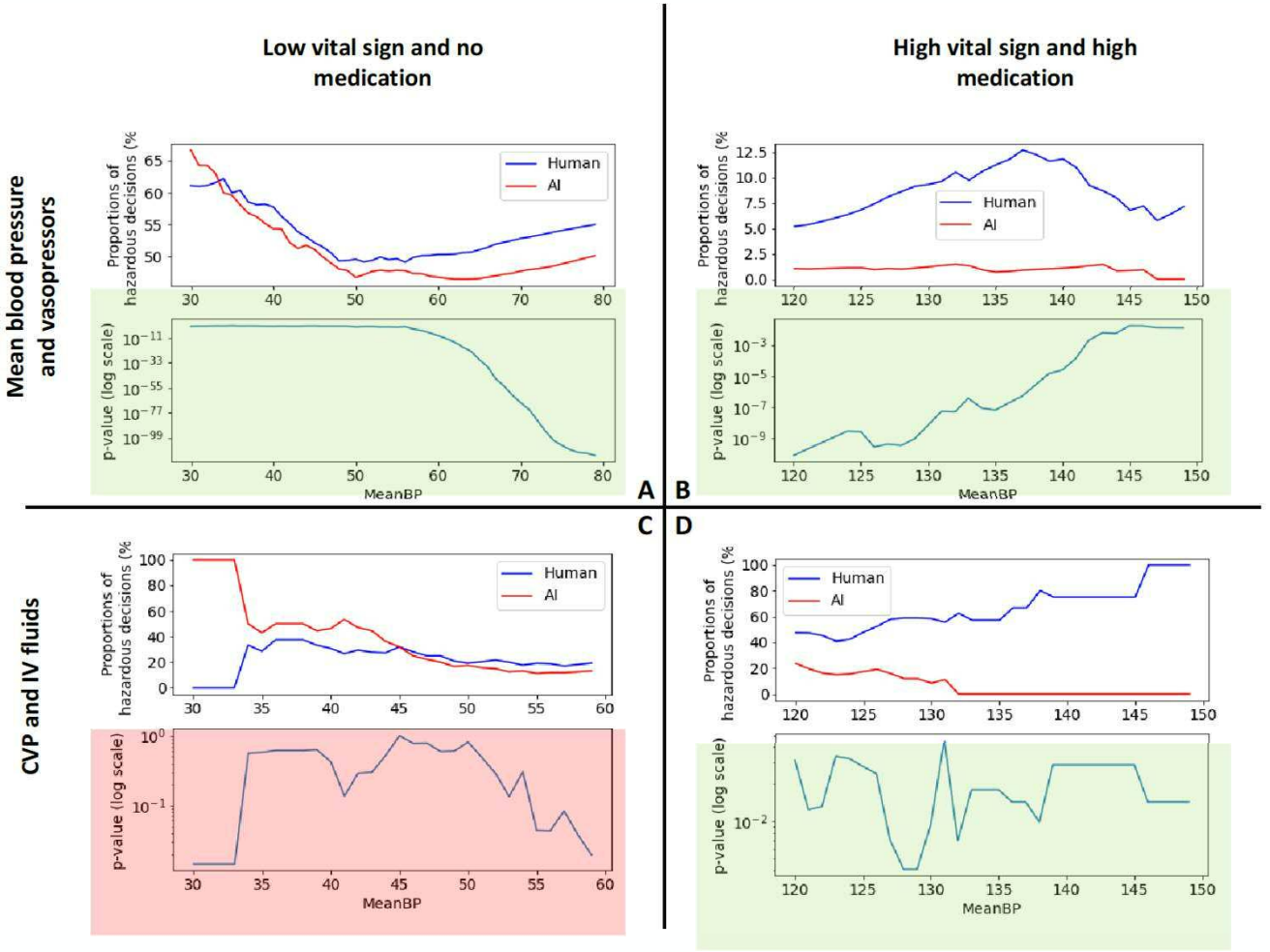
4 Hazardous scenarios

	Underdosing	Overdosing
Vasopressors	Scenario A	Scenario B
IV Fluids	Scenario C	Scenario D

Rule-based safety assessment

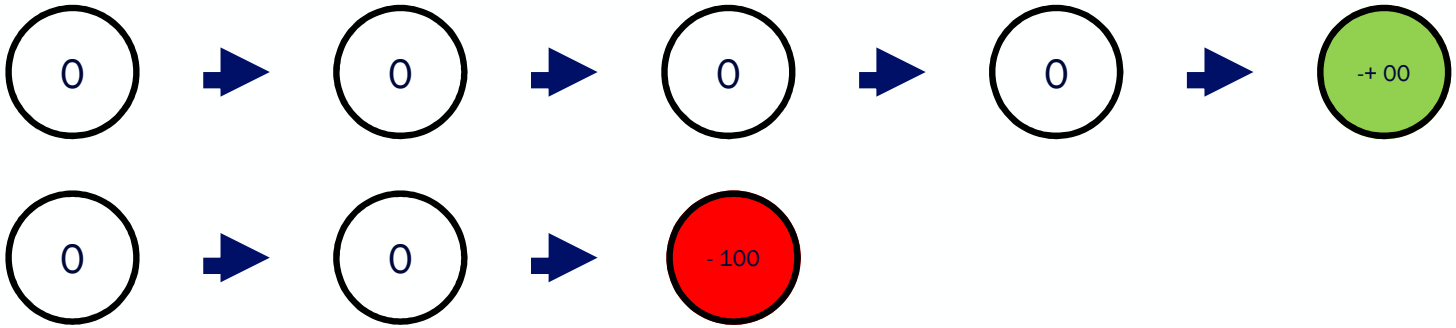


Rule-based safety assessment

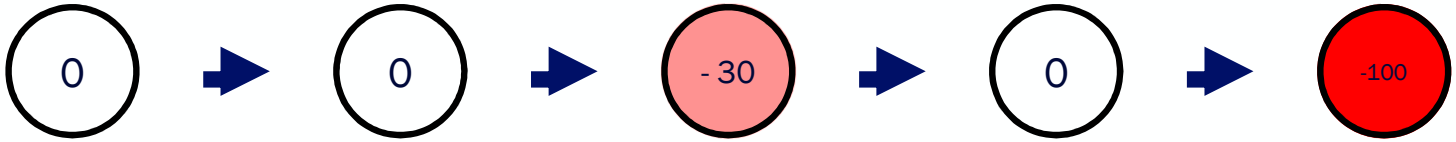


Feeding back into the model: intermediate rewards

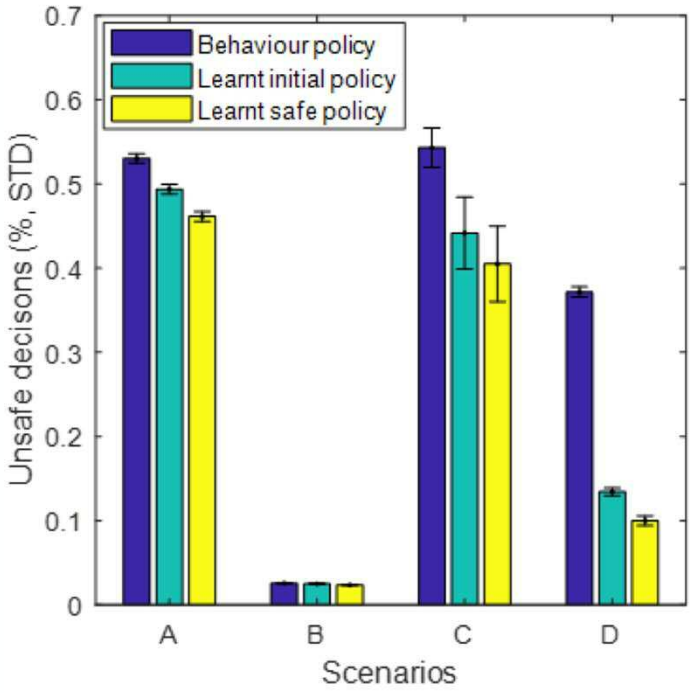
- Original reward for 90 days survival only:

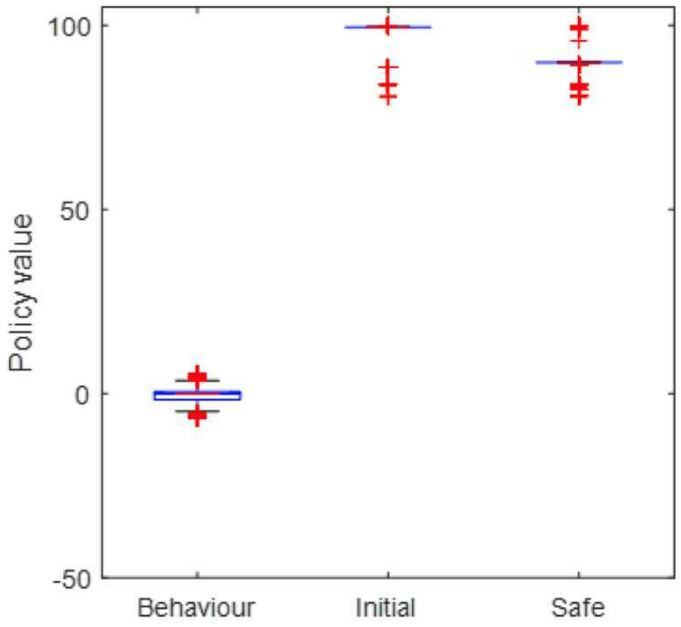


- Intermediate penalty for hazardous recommendatioans



The AI is safer on retrospective data and can even be improved

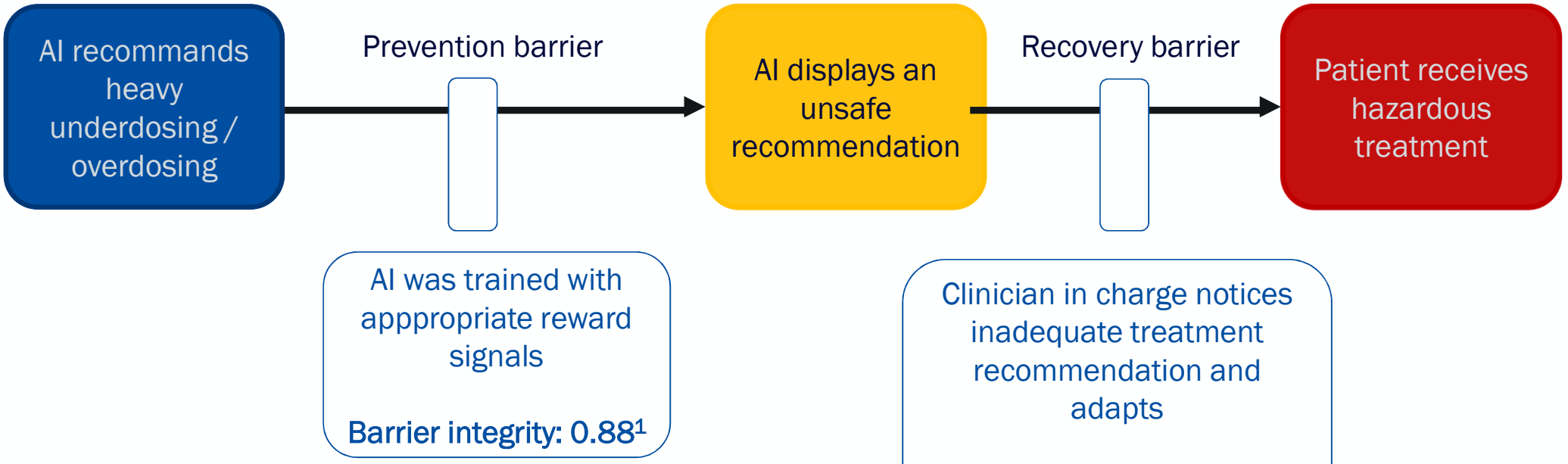




Festor et. Komorowski, “Assuring the safety of AI-based clinical decision support systems: a case study of the AI Clinician for sepsis treatment”, BMJ HCI (in production)

Underdosing/overdosing hazard analysis diagram

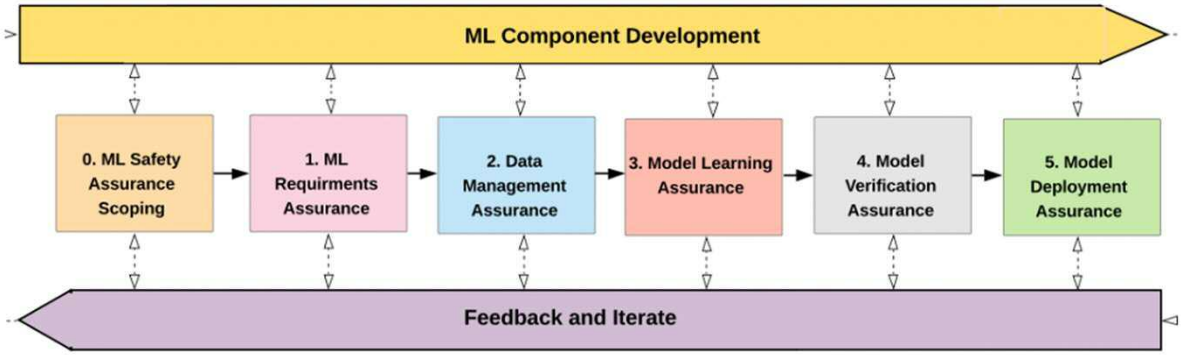
- Given a patient in one of the 4 hazardous scenarios (16% of our MIMIC cohort)



[1]Festor et. Komorowski, "Assuring the safety of AI-based clinical decision support systems: a case study of the AI Clinician for sepsis treatment", BMJ HCI (in production)

Safety assurance: discussion

- Safety assessment is iterative



- Limitation

"Safe" defined by rules → Safe decisions are known → No need for an AI

Hawkins, Richard, et al. "Guidance on the assurance of machine learning in autonomous systems (AMLAS)." *arXiv preprint arXiv:2102.01564* (2021).

Phd Interests: ethical and philosophical implications

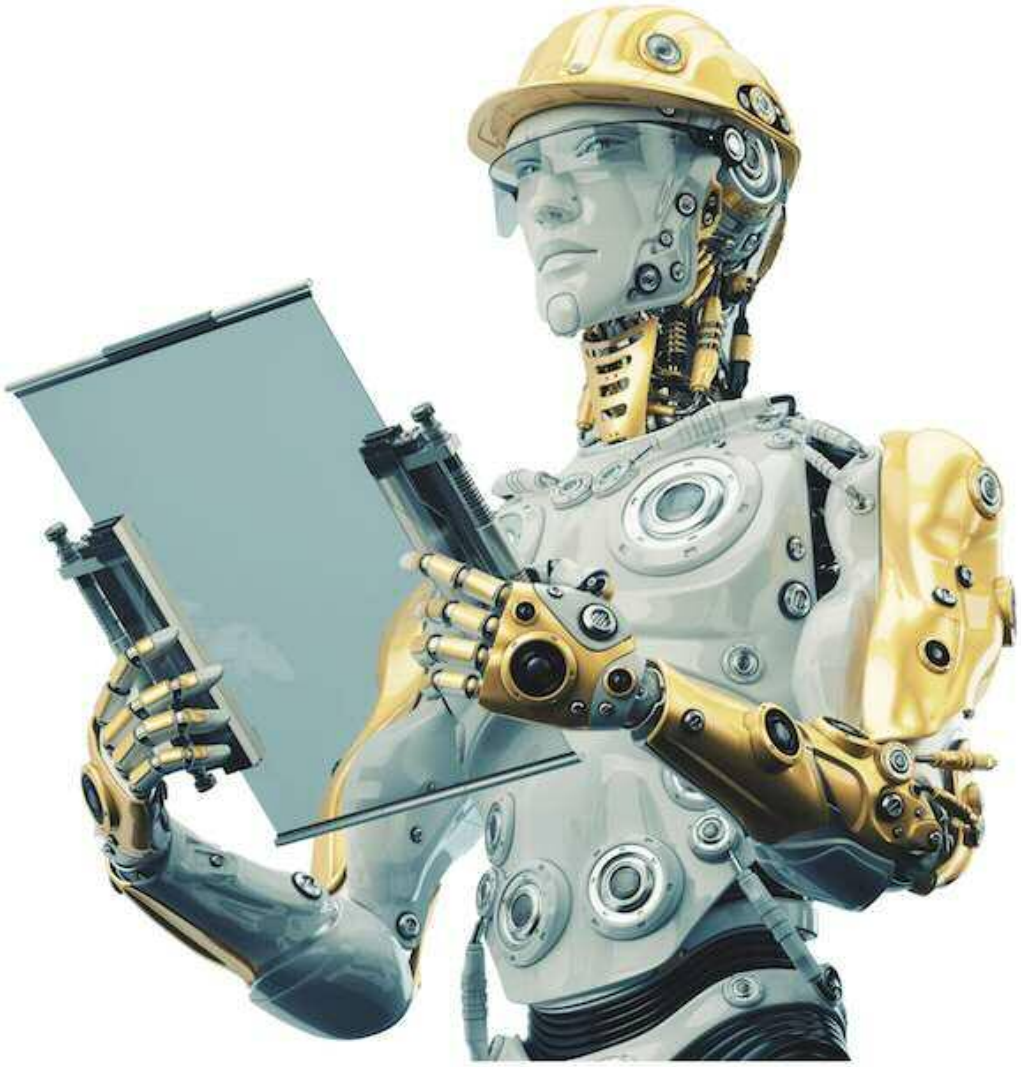
Levels of Autonomy & Safety Assurance for AI-based Clinical Decision Systems

Paul Festor[1,2], Ibrahim Habli[3,4], Yan Jia[3,4], Anthony Gordon[5], A. Aldo Faisal[1,2,6], and Matthieu Komorowski[1,5,6]

[1] UKRI Centre for Doctoral Training in AI for Healthcare, Imperial College London, UK
[2] Dept. of Computing, Dept. of Computing, Imperial College London, UK
[3] Assuring Autonomy Intl. Programme, University of York, UK
[4] Dept. of Computer Science, University of York, UK
[5] Dept. of Surgery & Cancer, Imperial College London, UK
[6] Dept. of Bioengineering, Dept. of Computing, Imperial College London, UK
[7] Corresponding author: m.komoroswki14@imperial.ac.uk

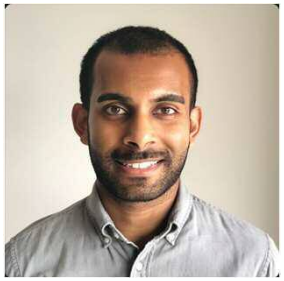
Conclusion

- Promising results in AI for treatment recommendation **BUT** large research-bedside gap
- Can leverage human knowledge to make AI safer
- Safety assessment is iterative
- Other interesting directions: uncertainty-aware, explainable systems & real world trials



Our team

Paul Festor



Myura Nagrendran

Anthony Gordon

Aldo Faisal

Matthieu Komorowski

Yan Jia

Ibrahim Habli

Imperial College London

UK Research and Innovation

ASSURING AUTONOMY INTERNATIONAL PROGRAMME

UNIVERSITY of York

SAFETY ASSESSMENT OF THE AI CLINICIAN FOR SEPSIS TREATMENT: AN ITERATIVE APPROACH

05/07/2022

Paul Festor - Future of Health 2022
Project subervised by A. Aldo Faisal & Matthieu Komorowski

References

McKinney, Scott Mayer, et al. "International evaluation of an AI system for breast cancer screening." *Nature* 577.7788 (2020): 89-94.

Hyland, Stephanie L., et al. "Early prediction of circulatory failure in the intensive care unit using machine learning." *Nature medicine* 26.3 (2020): 364-373.

Komorowski, Matthieu, et al. "The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care." *Nature medicine* 24.11 (2018): 1716-1720.

Johnson, Alistair EW, et al. "MIMIC-III, a freely accessible critical care database." *Scientific data* 3.1 (2016): 1-9.

Pollard, Tom J., et al. "The eICU Collaborative Research Database, a freely available multi-center database for critical care research." *Scientific data* 5.1 (2018): 1-13.

Thoral, Patrick J., et al. "Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example." *Critical care medicine* 49.6 (2021): e563.

van de Sande, Davy, et al. "Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit." *Intensive care medicine* (2021): 1-11.

York Health Economics Consortium, "The cost of sepsis care in the UK" (2017)

World Health Organization. "Global report on the epidemiology and burden of sepsis: Current evidence, identifying gaps and future directions." (2020).

Asaadi et. Petroff. Dynamic assurance cases: a pathway to trusted autonomy. *Computer* **53**, 35–46 (2020).

Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

Hawkins, Richard, et al. "Guidance on the assurance of machine learning in autonomous systems (AMLAS)." *arXiv preprint arXiv:2102.01564* (2021).

Published work

Festor, Paul, et al. "Levels of Autonomy and Safety Assurance for AI-Based Clinical Decision Systems." *International Conference on Computer Safety, Reliability, and Security*. Springer, Cham, 2021.

Festor, Paul, et al. "Enabling risk-aware Reinforcement Learning for medical interventions through uncertainty decomposition*" ICML Workshop on Interpretable Machine Learning for Healthcare*, 2021

Festor, Paul. et. "Assuring the safety of AI-based clinical decision support systems: a case study of the AI Clinician for sepsis treatment", BMJ Health & Care Informatics (in production)