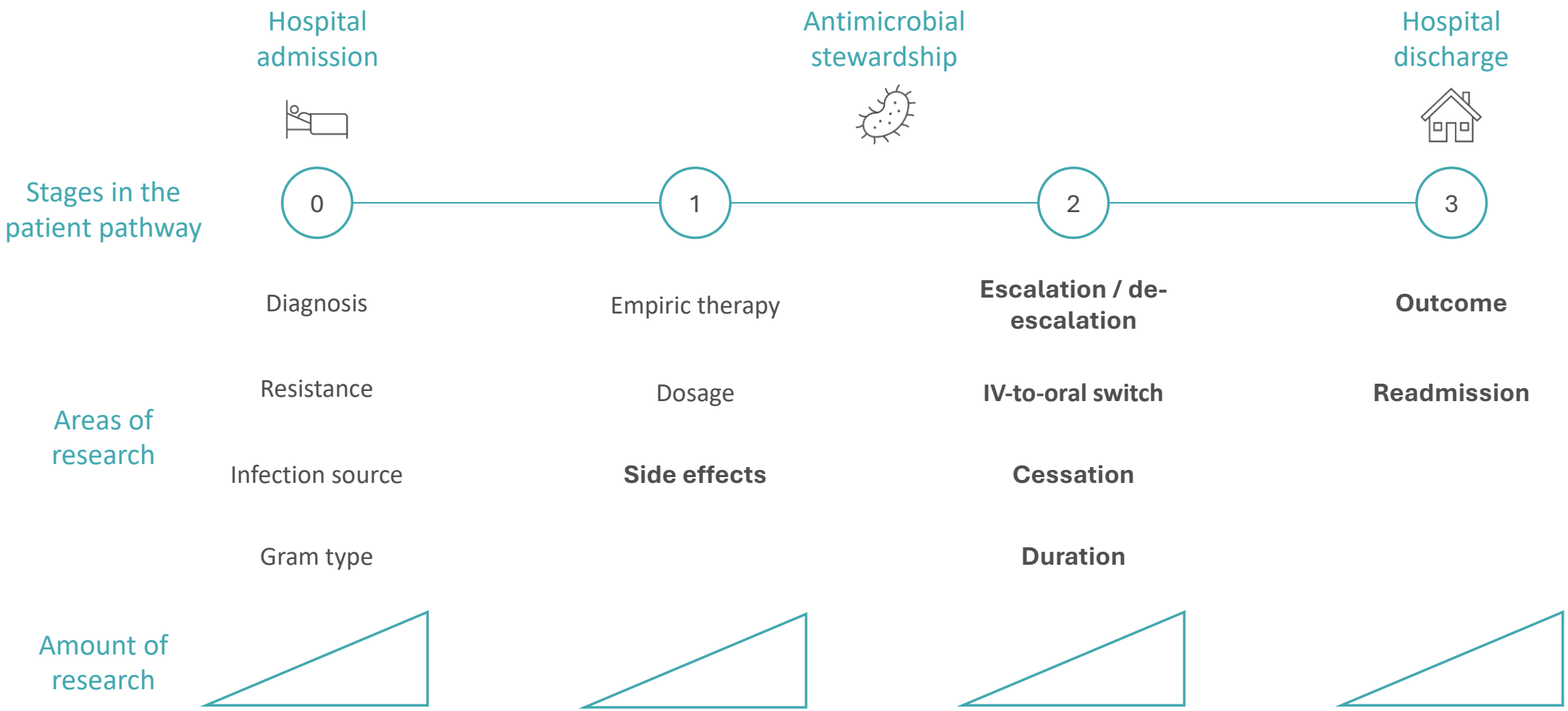
Prospective evaluation of a machine learning decision support system for intravenous-to-oral antibiotic switching

Dr William Bolton

ESCMID Global

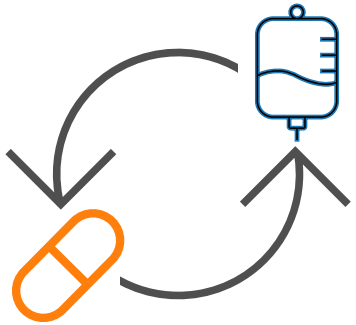
15th April 2025

Research into AI and data-driven approaches towards antimicrobial stewardship are lacking.



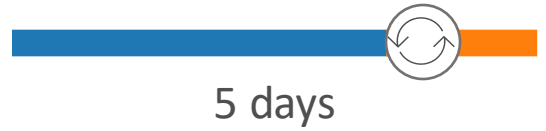
Oral antibiotics have numerous advantages, but switching from IV treatment is complex and under-researched.

make the switch
FROM IV TO ORAL ANTIBIOTICS



One key challenge of stewardship is **determining when to switch** antibiotics from **IV-to-oral administration**

Oral therapy are often **non-inferior** to IV with **fewer side effects**, decreased **nursing workload**, lower **costs**, reduced **climate impact** and improved **patient comfort**

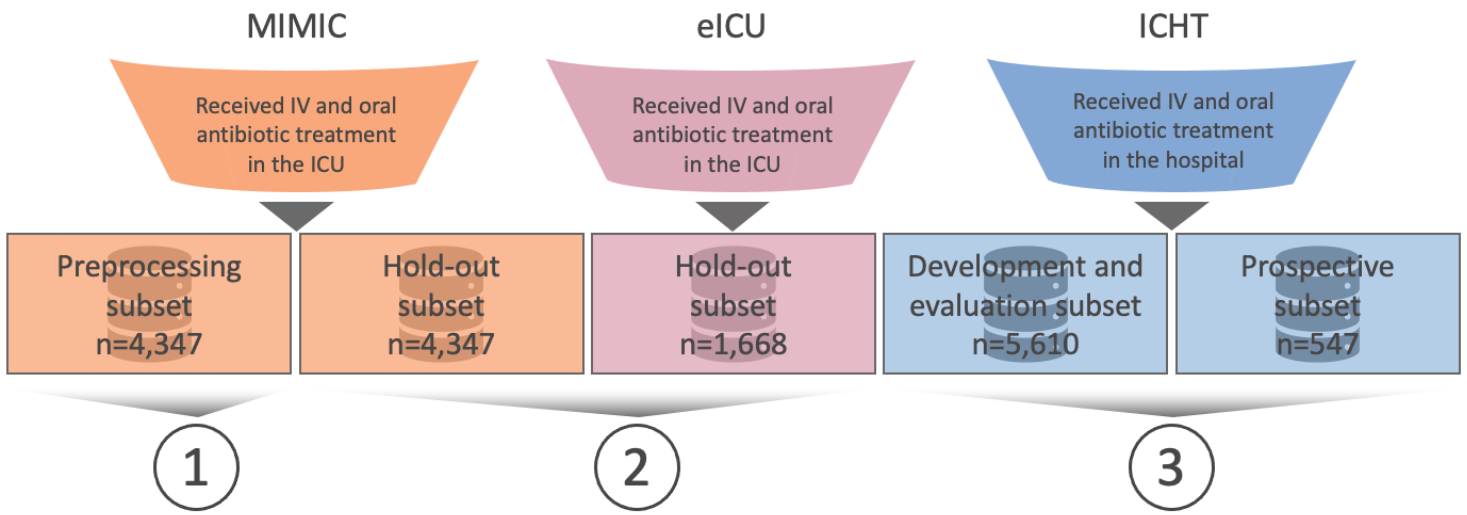


There is a **poor understanding** of the factors that facilitate or inhibit an individual from receiving oral therapy

Hypothesis

A **machine learning** model using **routinely collected clinical parameters** could predict whether a patient could be **suitable for switching** from IV-to-oral antibiotics on **any given day**

Machine learning models were trained to predict a patient's route of administration and evaluated across numerous datasets.



UKHSA IVOS Criteria

Two feature sets

Optimised models

Evaluation versus baseline

Subgroup analysis

IV treatment length

Fairness

Interpretability

Developed models that also incorporated demographics and co-morbidities

Evaluation including subgroups and prospective dataset

Point prevalence survey

nature communications

Models achieve generalisable performance across a range of datasets and patient populations.

MIMIC

Metric	1ST threshold results	2nd threshold results	IVOS criteria baseline
AUROC	**0.78 (SD 0.02)**	0.69 (SD 0.03)	0.66
FPR	0.25 (SD 0.02)	**0.10 (SD 0.02)**	0.43

eICU

Metric	1st threshold results	2nd threshold results	IVOS criteria baseline
AUROC	**0.72 (SD 0.02)**	0.65 (SD 0.05)	0.55
FPR	0.24 (SD 0.04)	**0.05** (SD 0.02)	0.28

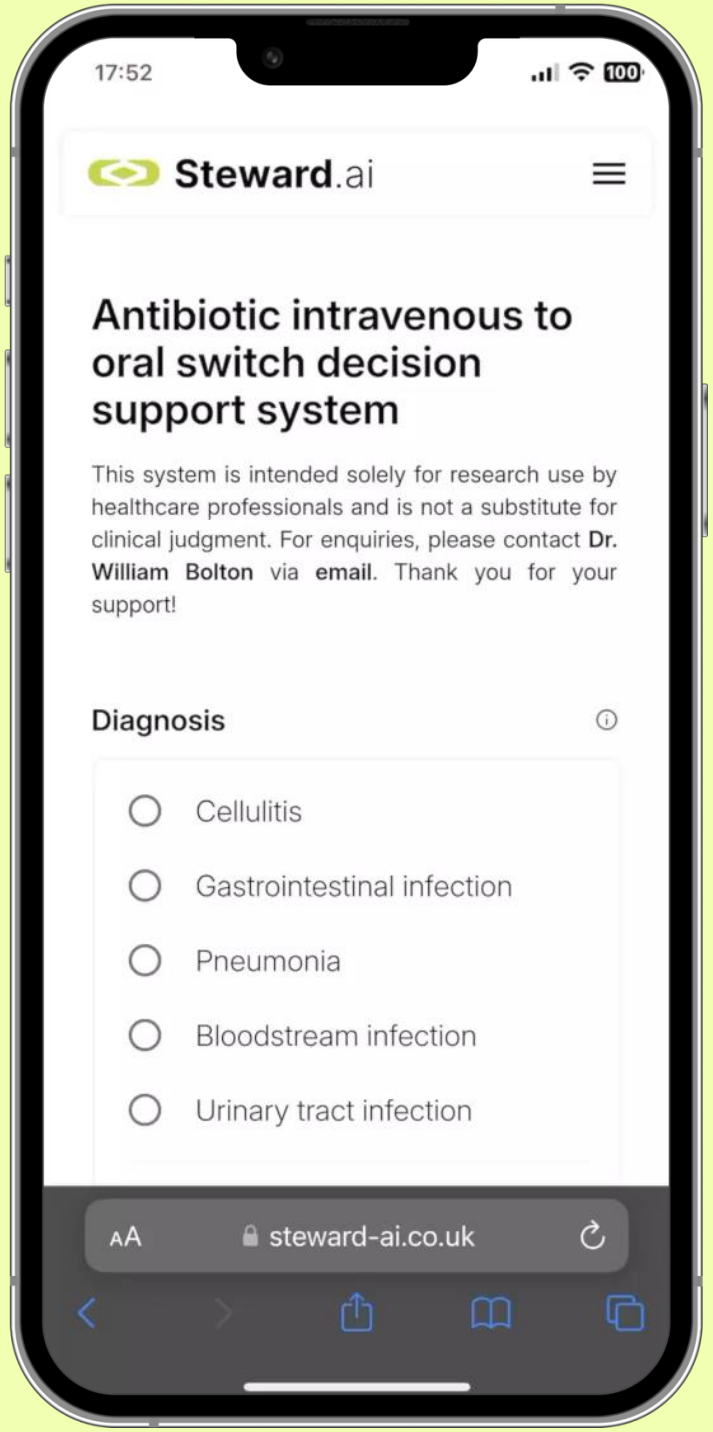
NHS Imperial College Healthcare NHS Trust

Metric	Retrospective dataset	Prospective dataset
AUROC	0.79 (SD 0.01)	0.77
FPR	0.21 (SD 0.03)	0.20

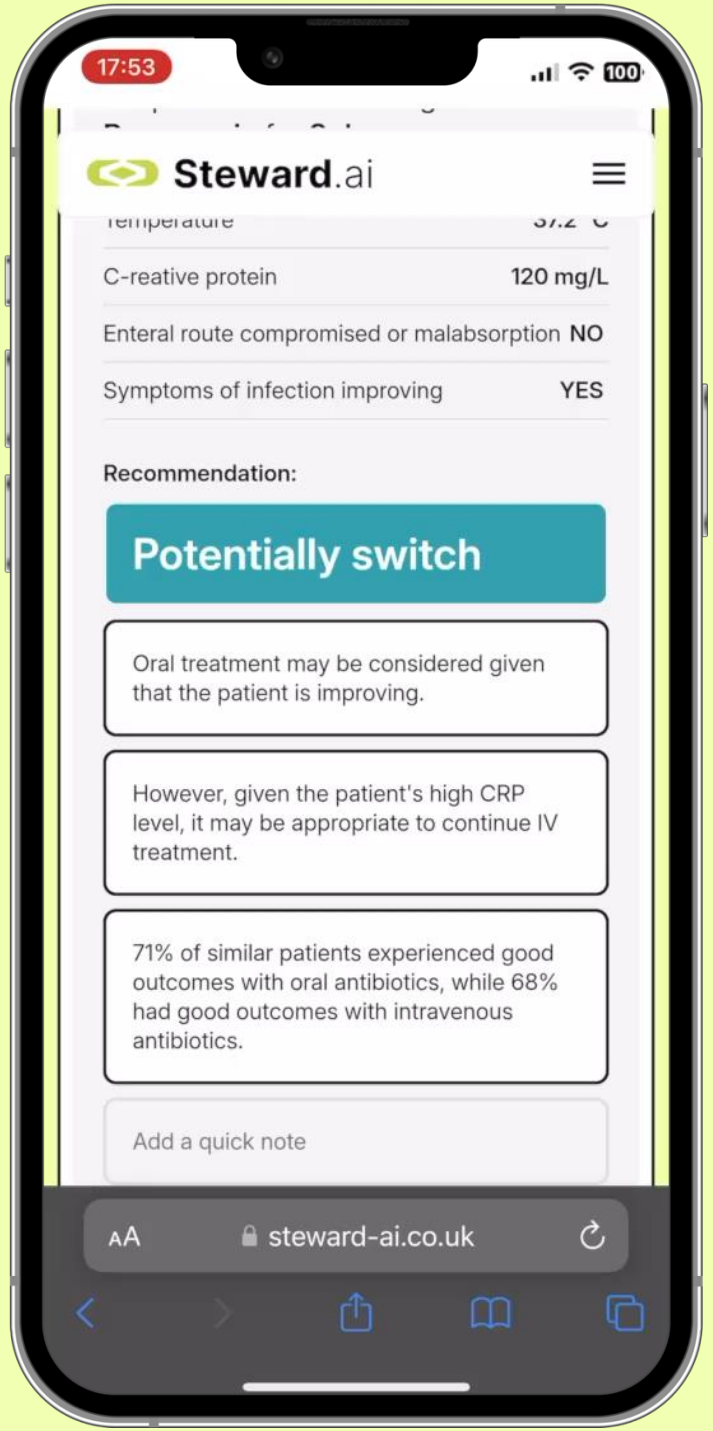
nature communications

Steward.ai

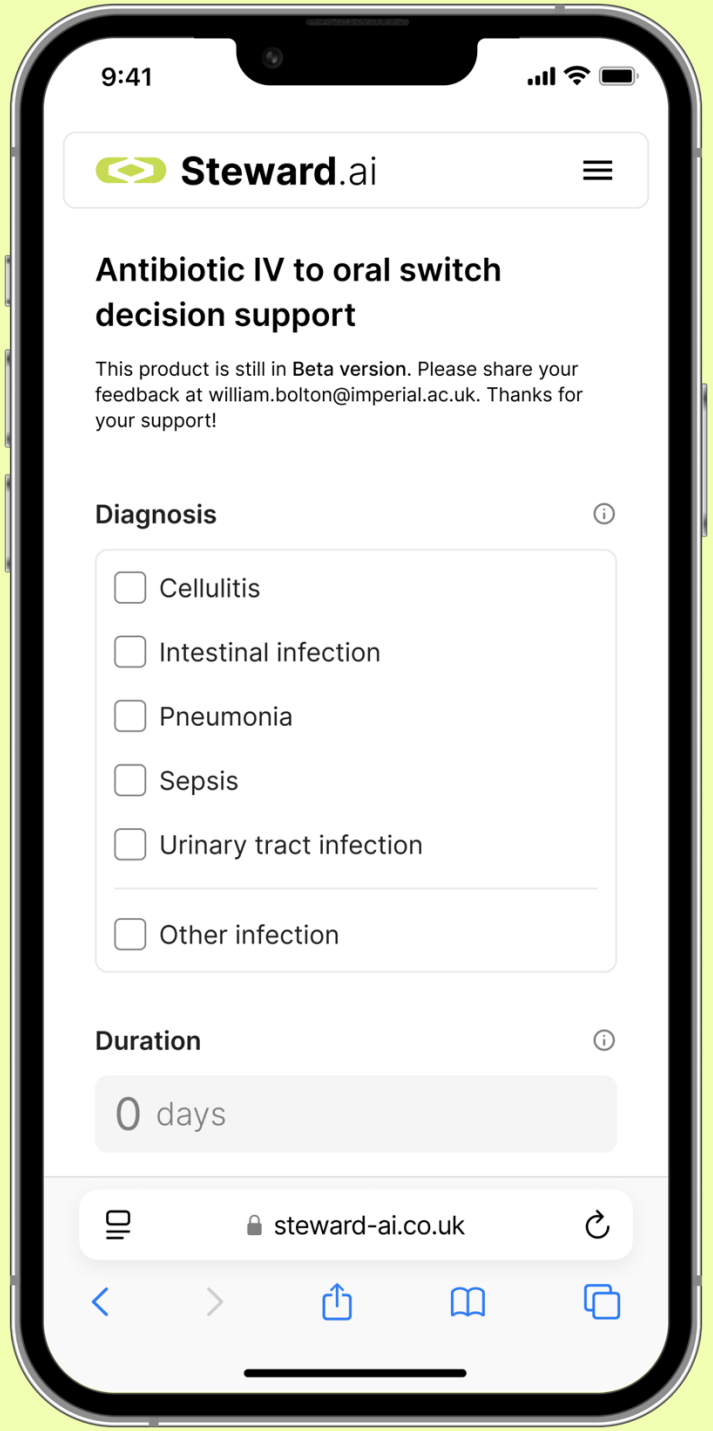
Empowering Clinicians Towards
Precision Antibiotic Therapy


17:52
Steward.ai
Antibiotic intravenous to oral switch decision support system
This system is intended solely for research use by healthcare professionals and is not a substitute for clinical judgment. For enquiries, please contact Dr. William Bolton via email. Thank you for your support!
Diagnosis
Cellulitis
Gastrointestinal infection
Pneumonia
Bloodstream infection
Urinary tract infection
AA
steward-ai.co.uk

Steward.ai

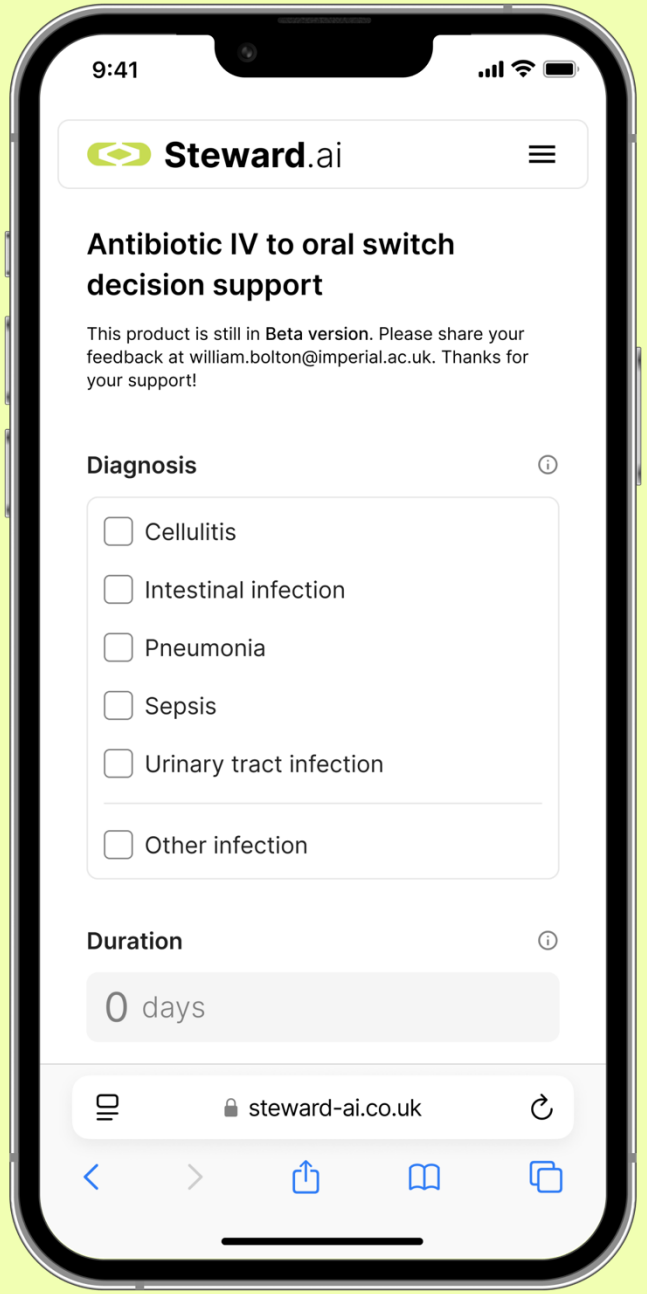

17:53
100
Steward.ai
C-reative protein 120 mg/L
Enteral route compromised or malabsorption NO
Symptoms of infection improving YES
Recommendation:
Potentially switch
Oral treatment may be considered given that the patient is improving.
However, given the patient's high CRP level, it may be appropriate to continue IV treatment.
71% of similar patients experienced good outcomes with oral antibiotics, while 68% had good outcomes with intravenous antibiotics.
Add a quick note
AA
steward-ai.co.uk

Steward.ai


9:41
Steward.ai
Antibiotic IV to oral switch decision support
This product is still in Beta version. Please share your feedback at william.bolton@imperial.ac.uk. Thanks for your support!
Diagnosis
Cellulitis
Intestinal infection
Pneumonia
Sepsis
Urinary tract infection
Other infection
Duration
0 days
steward-ai.co.uk

Steward.ai

- **Personalized** patient decisions
- Simple and **easy to use**
- Explainable, **safeguarded AI**
- Robust **clinical evaluation**
- Save **hospitals money**



Prospective evaluation performance of Steward.ai is promising with a low false positive rate.

Professor Mark Gilchrist

Bi-annual Antimicrobial **Point Prevalence Survey** (PPS) conducted bt the Infection Pharmacy Team at ICHT

NHS Imperial College Healthcare NHS Trust

24 Patients receiving IVOS relevant antibiotics (e.g., co-amoxiclav)

68 days of antibiotic treatment

Metric	Prior model results	**Steward**.ai results Excluding potentially switch prediction	**Steward**.ai results Including potentially switch prediction
AUROC	0.68		
FPR	0.28		
Accuracy	0.70		
Precision	-		
Recall	-		
F1 score	-		
Weighted Partial Credit Accuracy	-		
Weighted Cohen's Kappa	-		

Prospective evaluation performance of Steward.ai is promising with a low false positive rate.

Professor Mark Gilchrist

Bi-annual Antimicrobial **Point Prevalence Survey** (PPS) conducted bt the Infection Pharmacy Team at ICHT

NHS Imperial College Healthcare NHS Trust

24 Patients receiving IVOS relevant antibiotics (e.g., co-amoxiclav)

68 days of antibiotic treatment

Metric	Prior model results	**Steward**.ai results Excluding potentially switch prediction	**Steward**.ai results Including potentially switch prediction
AUROC	0.68	**0.76**	-
FPR	0.28	**0.06**	-
Accuracy	0.70	**0.81**	-
Precision	-	0.85	-
Recall	-	0.58	-
F1 score	-	0.69	-
Weighted Partial Credit Accuracy	-	-	0.74
Weighted Cohen's Kappa	-	-	0.42

Prospective evaluation performance of Steward.ai is promising with temporal alignment with pharmacists.

Professor Mark Gilchrist

Bi-annual Antimicrobial **Point Prevalence Survey** (PPS) conducted bt the Infection Pharmacy Team at ICHT

NHS Imperial College Healthcare NHS Trust

24 Patients receiving IVOS relevant antibiotics (e.g., co-amoxiclav)

68 days of antibiotic treatment

Temporal difference relative to pharmacists	Prior model results	**Steward**.ai results
Before	37.5%	
Same day	33.5%	
After	30%	

Prospective evaluation performance of Steward.ai is promising with temporal alignment with pharmacists.

Professor Mark Gilchrist

Bi-annual Antimicrobial **Point Prevalence Survey** (PPS) conducted bt the Infection Pharmacy Team at ICHT

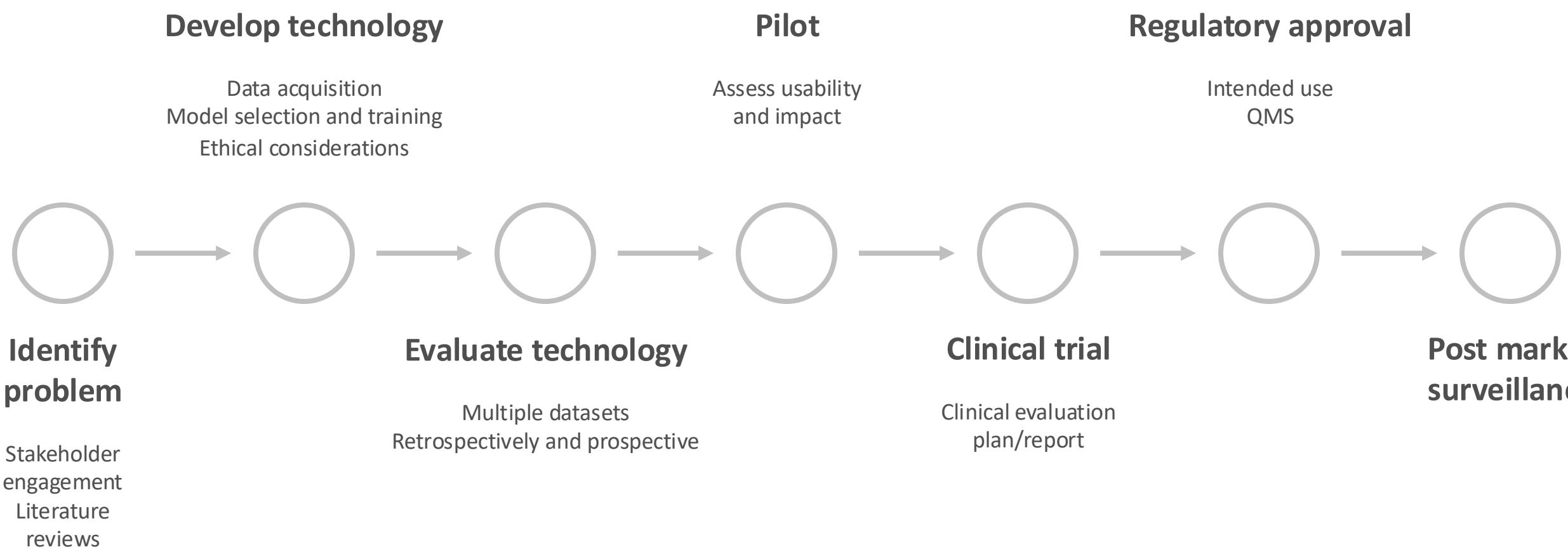
NHS Imperial College Healthcare NHS Trust

24 Patients receiving IVOS relevant antibiotics (e.g., co-amoxiclav)

68 days of antibiotic treatment

Temporal difference relative to pharmacists	Prior model results	**Steward**.ai results
Before	37.5%	**8%**
Same day	33.5%	**58%**
After	**30%**	33%

The roadmap for AI clinical decision support systems.



We are beginning a real-world pilot at Imperial and are looking to partner with other healthcare and MedTech organisations.

Developed technology

Web application, research published in leading journals

nature communications · nature machine intelligence · AAAI · ICLR · ESCMID

Piloted

Piloted in simulated environment, real-world clinical pilot beginning

NHS Imperial College Healthcare NHS Trust

Professor Mark Gilchrist

Regulatory approval

Intended use
QMS

2023 → 2023 → 2024 → 2025 → 2026 → → →

Problem identified

IV-to-oral antimicrobial switching

Evaluated technology

Multiple datasets and prospective point prevenance survey

NHS Imperial College Healthcare NHS Trust · AI SuperConnector · conceptionx PhD DEEP TECH STARTUPS

Clinical trial

We are looking to establish partnerships with other hospitals and MedTech businesses for commercialization

Post market surveillance

Thank you for the support.

Dr Tim Rawson

Professor Pantelis Georgiou

Professor Alison Holmes

Professor Mark Gilchrist

Richard Wilson

Charn Sangvirojkul

James Skilton

Imperial College London

NHS Imperial College Healthcare NHS Trust

FHealth

camo centre for antimicrobial optimisation

AI SuperConnector

UKRI Research England

Prospective evaluation of a machine learning decision support system for intravenous-to-oral antibiotic switching

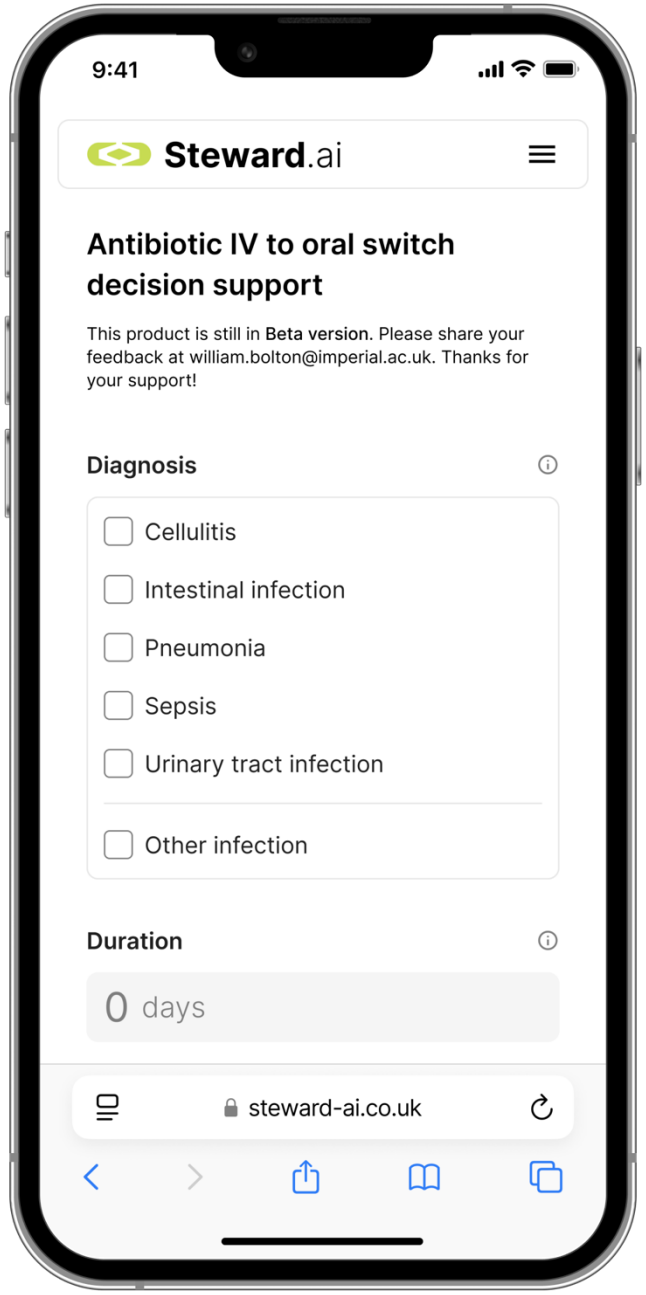
Dr William Bolton

william.bolton@imperial.ac.uk

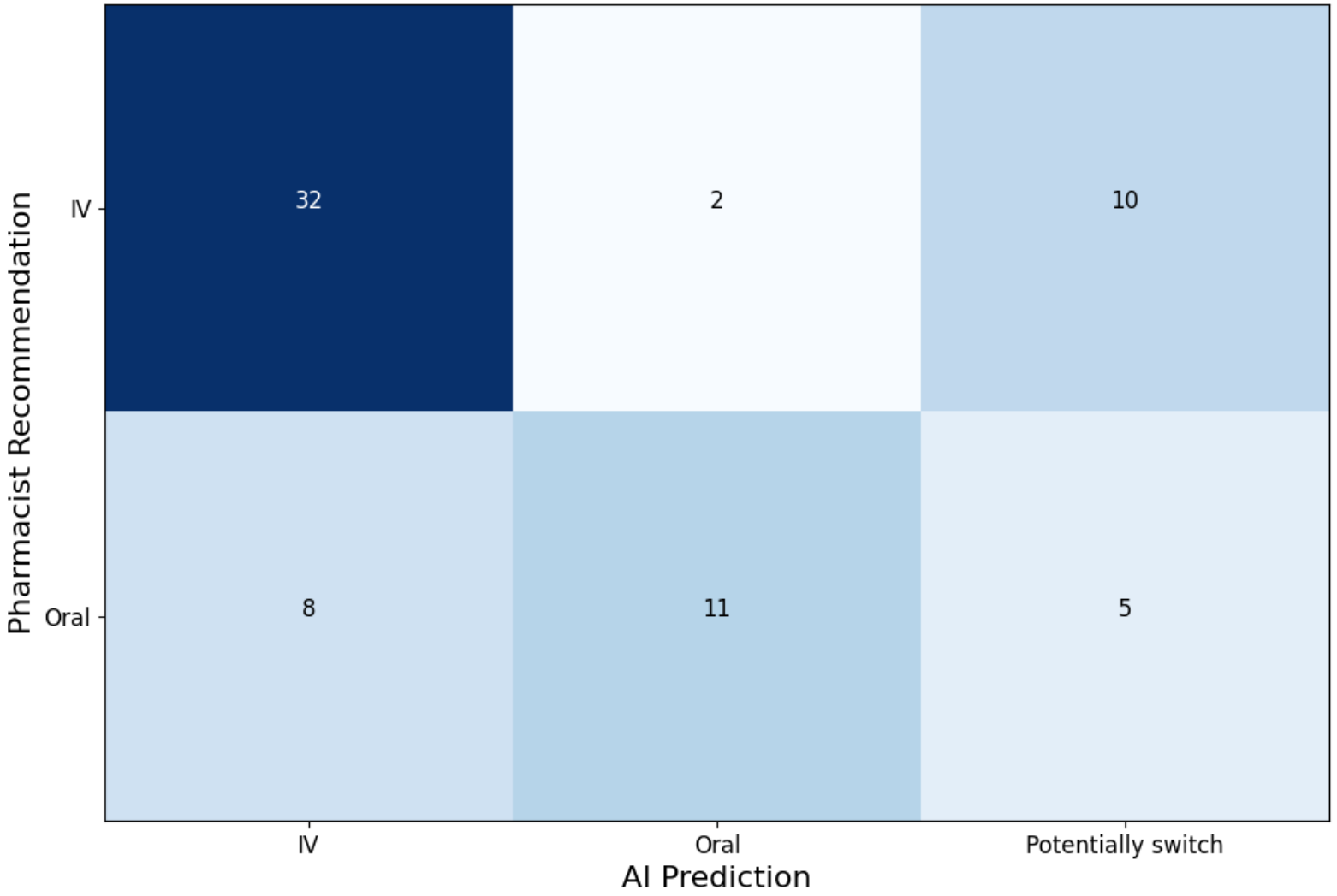
ESCMID Global

15th April 2025

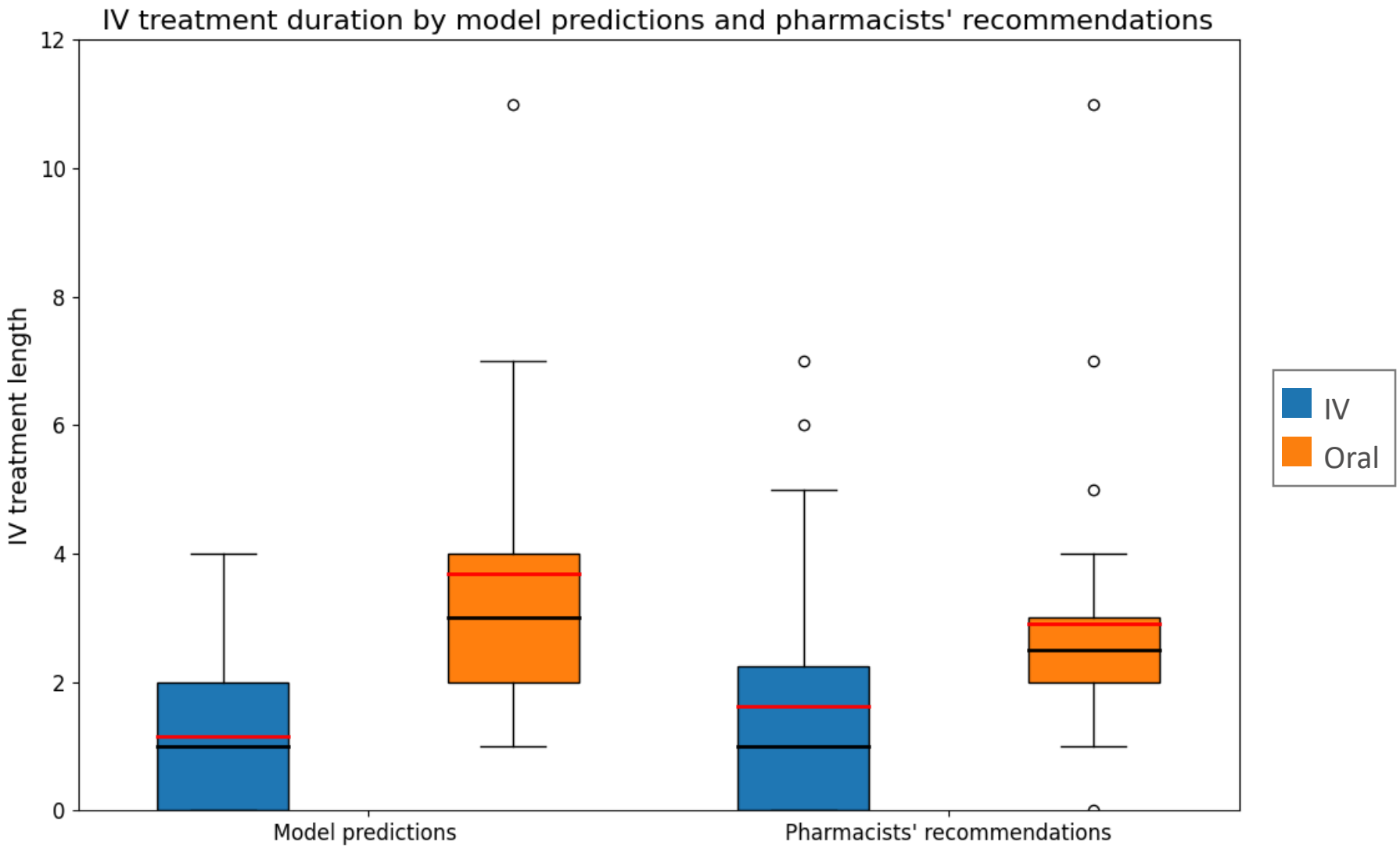
Try it out for yourself!



Prospective evaluation performance of Steward.ai is promising with a low false positive rate.



Prospective evaluation performance of Steward.ai is promising.



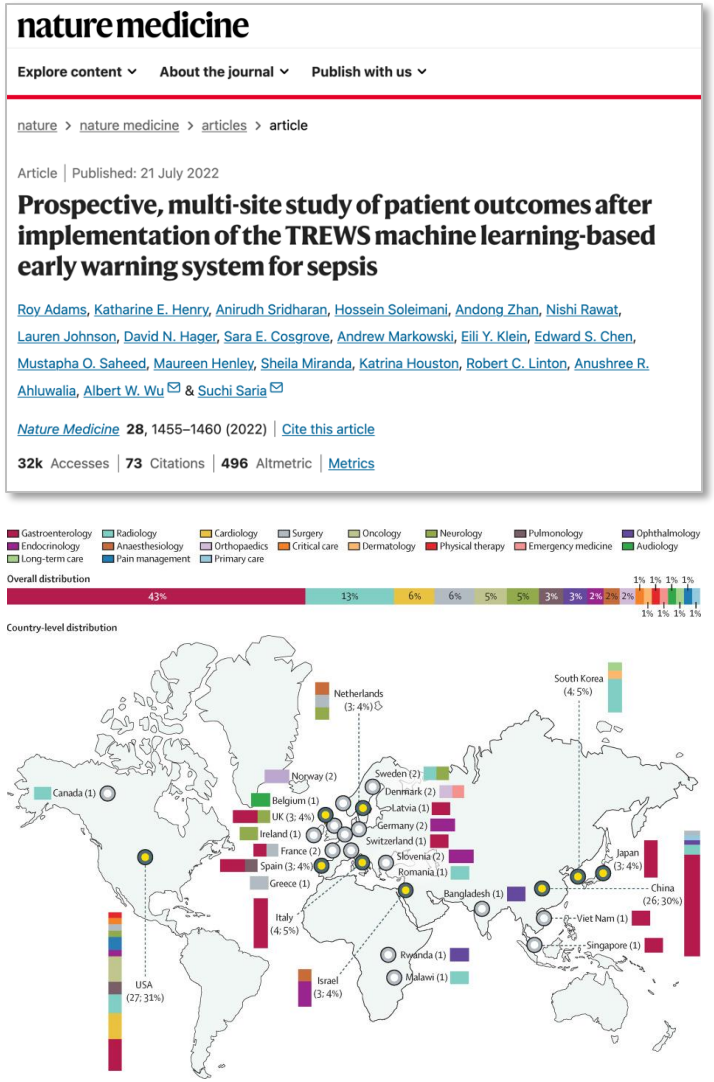
Using AI to optimize antimicrobial prescribing raises important ethical questions.

How can a **moral balance** be obtained between the needs of an **individual** patient and those of **wider and future society**?

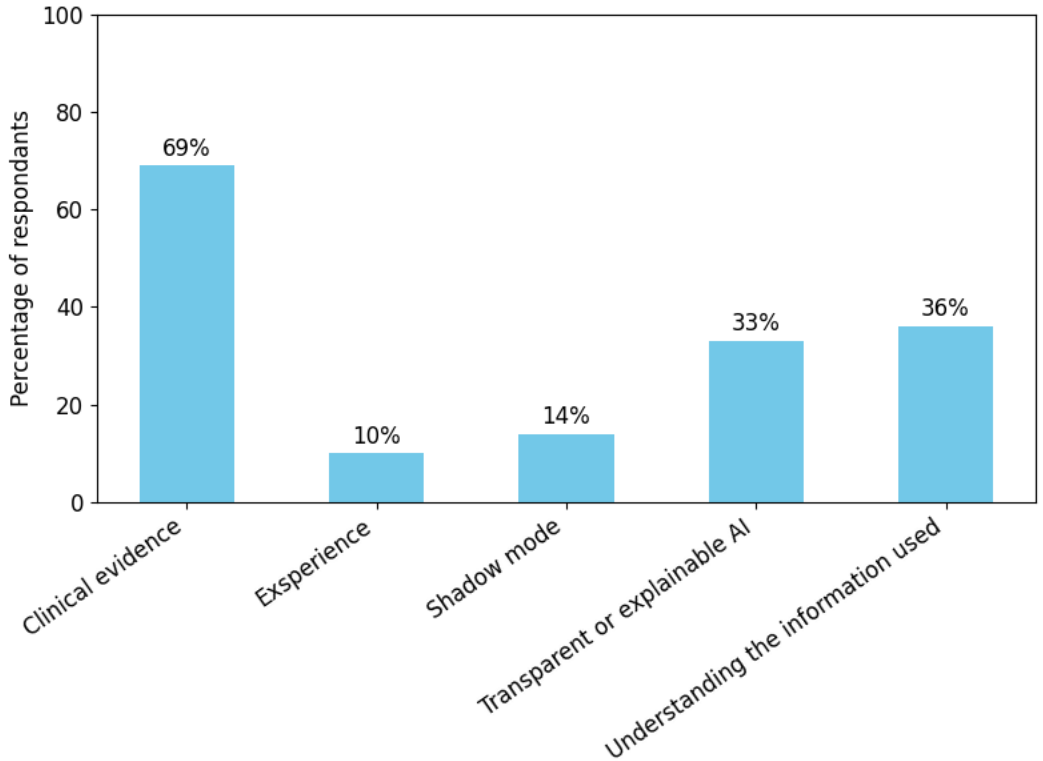
nature machine intelligence

Variables	Description	Exemplar of starting antimicrobial treatment	Corresponding ad-hoc utility value
Intensity	How strong is the pleasure?	Treating a relevant infection with antimicrobials has the potential to save that person's life	Highly positive utility
Duration	How long will the pleasure last?	Any extension of life is immeasurable while it is reasonable AMR will continue in the near-term future	Positive utility
Certainty or uncertainty	How likely or unlikely is it that the pleasure will occur?	Limited information often means treatment may or may not be helpful and there is always an inherent risk of developing AMR	Neutral utility, without more information
Propinquity	How soon will the pleasure occur?	Treatment can be effective immediately however the same is true for the evolution of AMR	Neutral utility, without more information
Fecundity	The likelihood of further sensations of the same kind	-	Unable to assign
Purity	The likelihood of not being followed by opposite sensations	-	Unable to assign
Extent	How many people will be affected?	Prescribing antimicrobials effects the patient and those close to them, while the development of AMR is a certainty and may affect everyone, causing significant suffering and mortality	Immense negative utility

Clinical evidence is necessary for trust but few clinical trials of AI in real clinical practice exist - especially in infectious diseases.



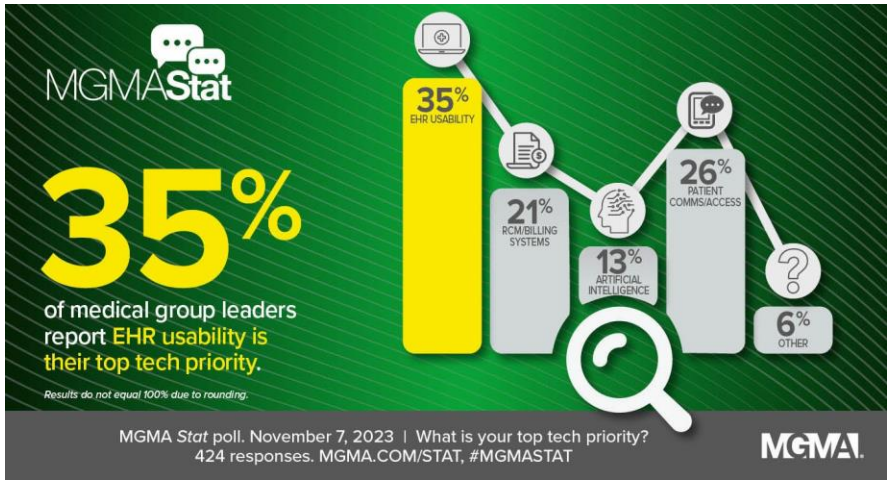
What factors determine the level of trust you have in an AI system?



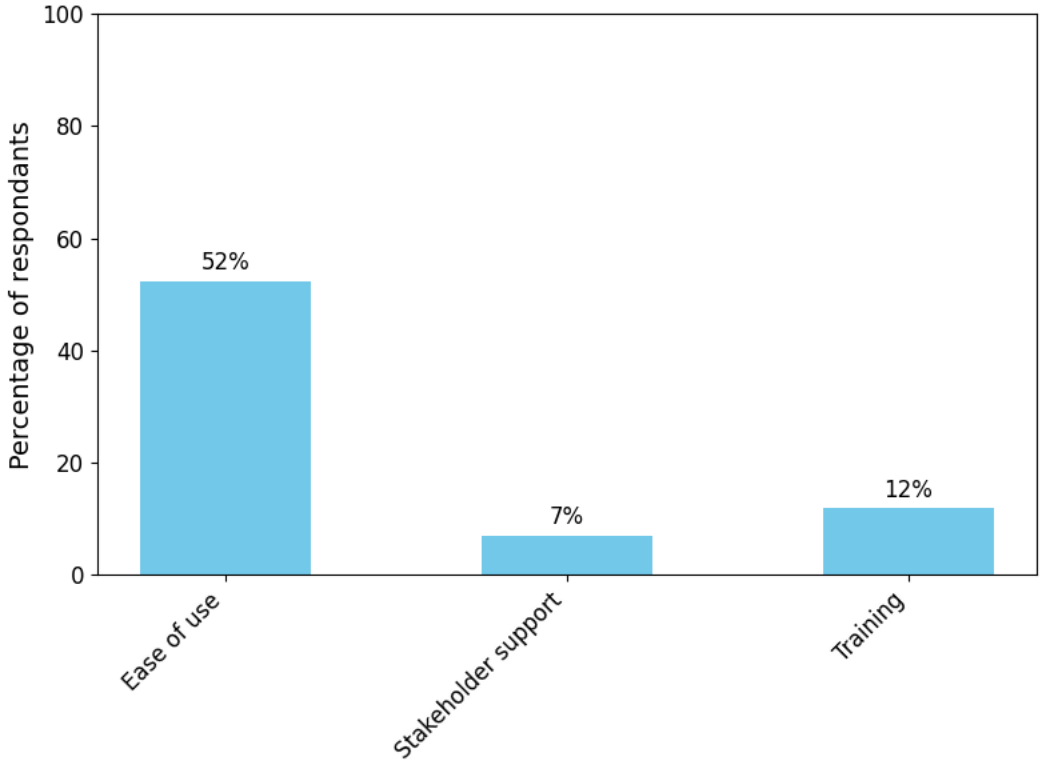
"Demonstration of utility and safety is absolutely critical."

Han R, Acosta JN, Shakeri Z, Ioannidis JP, Topol EJ, Rajpurkar P. Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. The Lancet Digital Health. 2024 May 1;6(5):e367-73.
Adams, R., Henry, K.E., Sridharan, A. et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. Nat Med 28, 1455–1460 (2022). https://doi.org/10.1038/s41591-022-01894-0

Usability is essential for real-world adoption.



What would convince you to consistently use an AI decision support system?



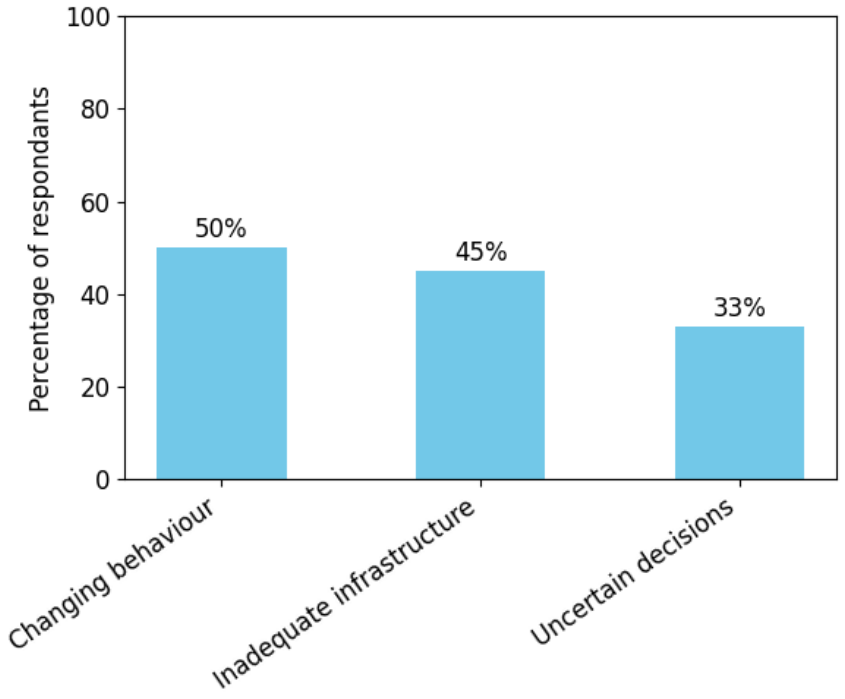
"I think the UI UX, and the whole design in terms of making it intuitive, is key."

https://www.mgma.com/mgma-stat/ehr-usability-patient-communications-billing-outrank-ai-as-top-tech-priorities

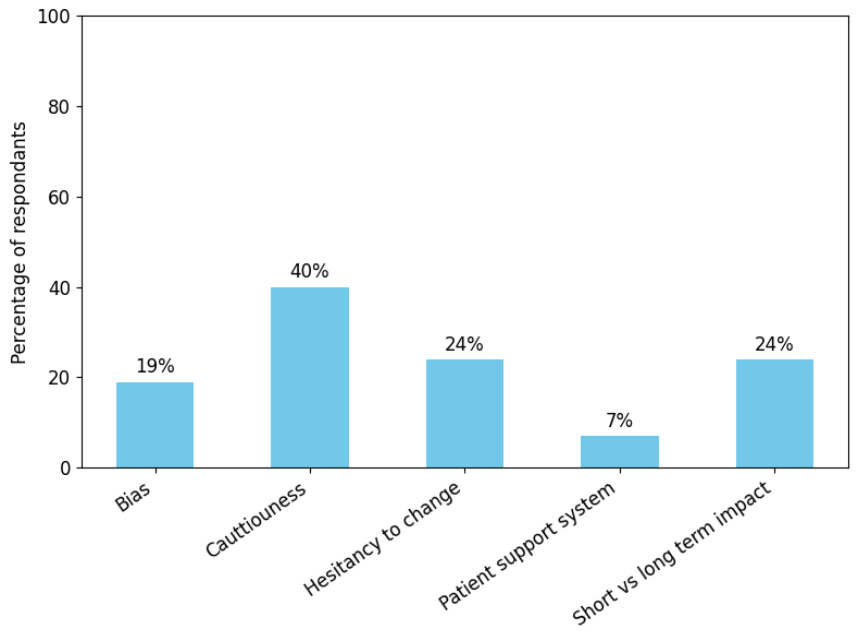
Maqbool, B. and Herold, S., 2023. Potential effectiveness and efficiency issues in usability evaluation within digital health: A systematic literature review. Journal of Systems and Software, p.111881.

Infrastructure, behaviour, culture and uncertainty pose challenges for AI systems in healthcare, particularly those focusing on AMR.

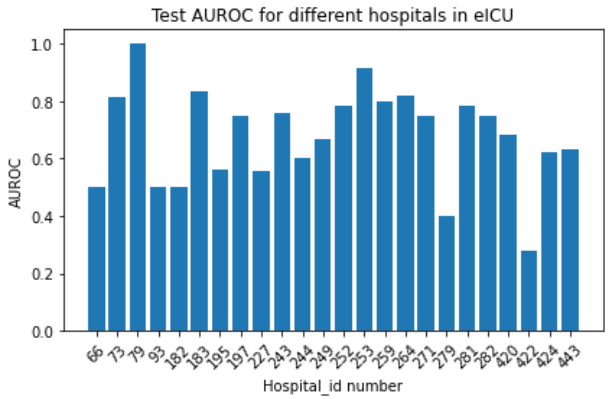
What is the biggest bottleneck in translating healthcare-focused AI research?



What factors influence antibiotic stewardship decisions?

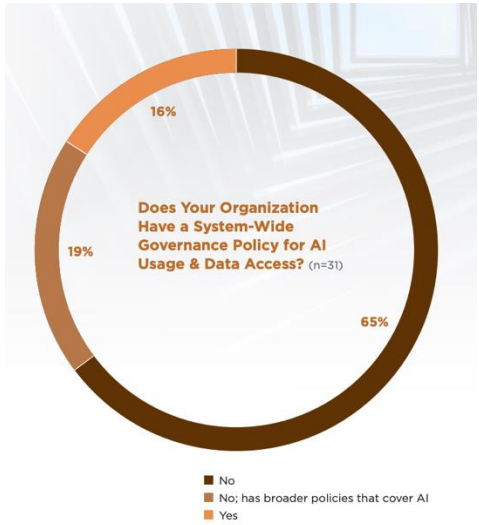


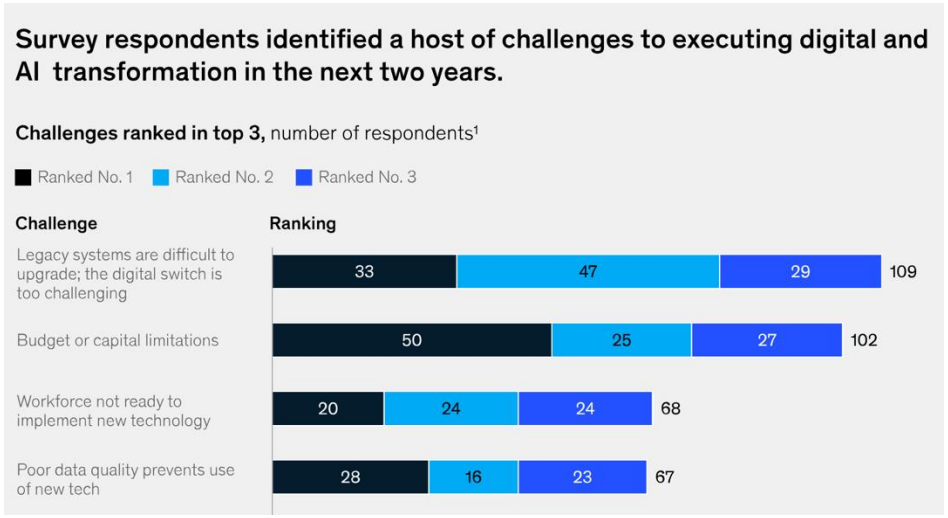
HUMAN BEHAVIOUR IS HETEROGENEOUS



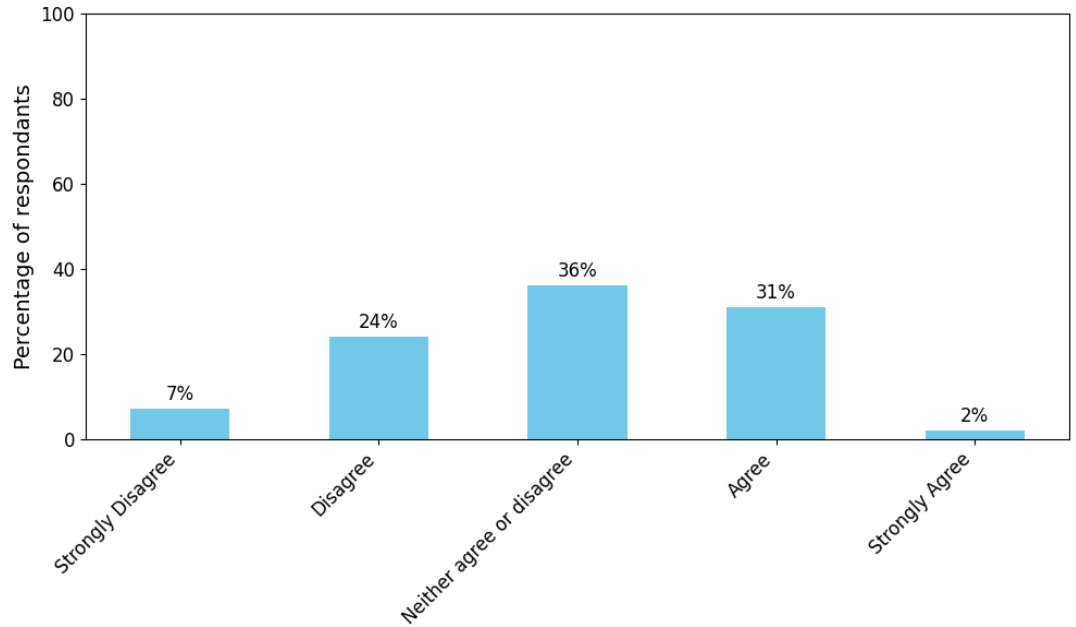
DATA QUALITY AND MISSINGNESS

Are hospitals ready for AI?





I think my healthcare institution has the necessary infrastructure to support this AI CDSS



31% of respondents in our study said they have not used any technology to support decision making

https://info.connectedmed.com/l/689353/2024-02-09/2lvknc/689353/1707510824kuJAqb0f/How_Health_Systems_Are_Navigating_The_Complexities_Of_AI_CCM_Reports.pdf?
https://www.mckinsey.com/industries/healthcare/our-insights/digital-transformation-health-systems-investment-priorities?utm_source=substack&utm_medium=email