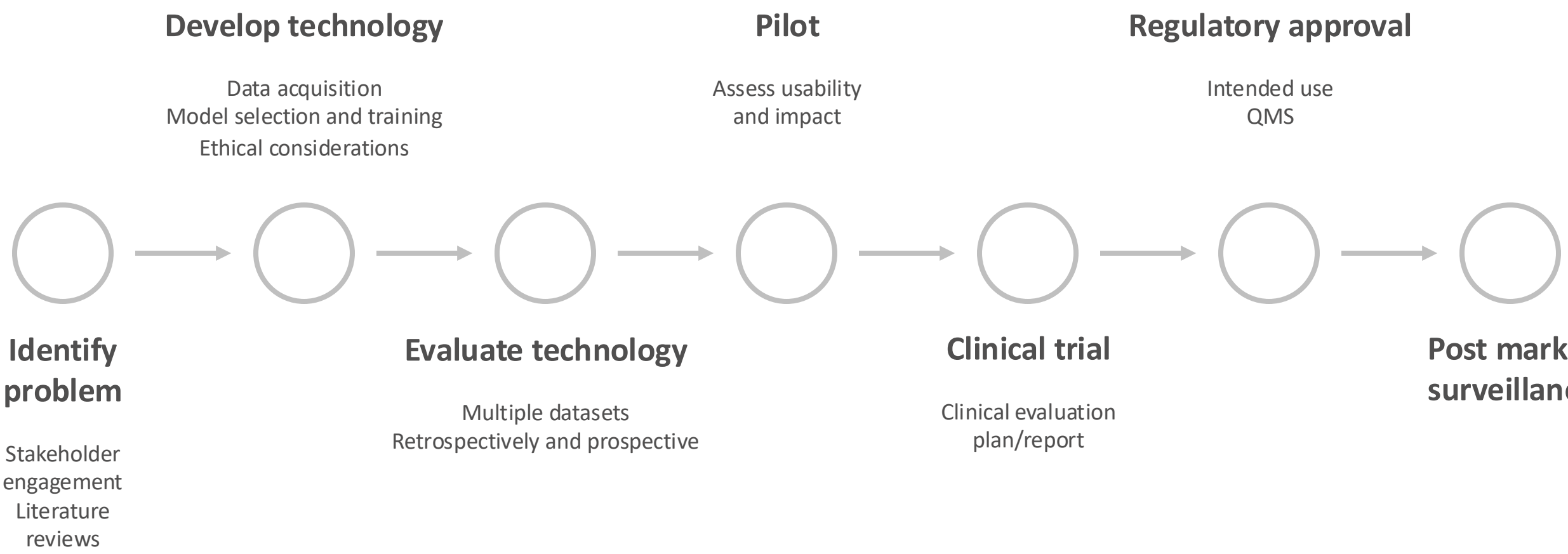
AI Decision Support for Antimicrobial Stewardship: Navigating Clinical Translation

Dr William Bolton

MedTech Links

24th March 2025

The roadmap for AI clinical decision support systems.



Do no harm: designing, evaluating and integrating AI in healthcare

nature medicine

Explore content ⌄ About the journal ⌄ Publish with us ⌄

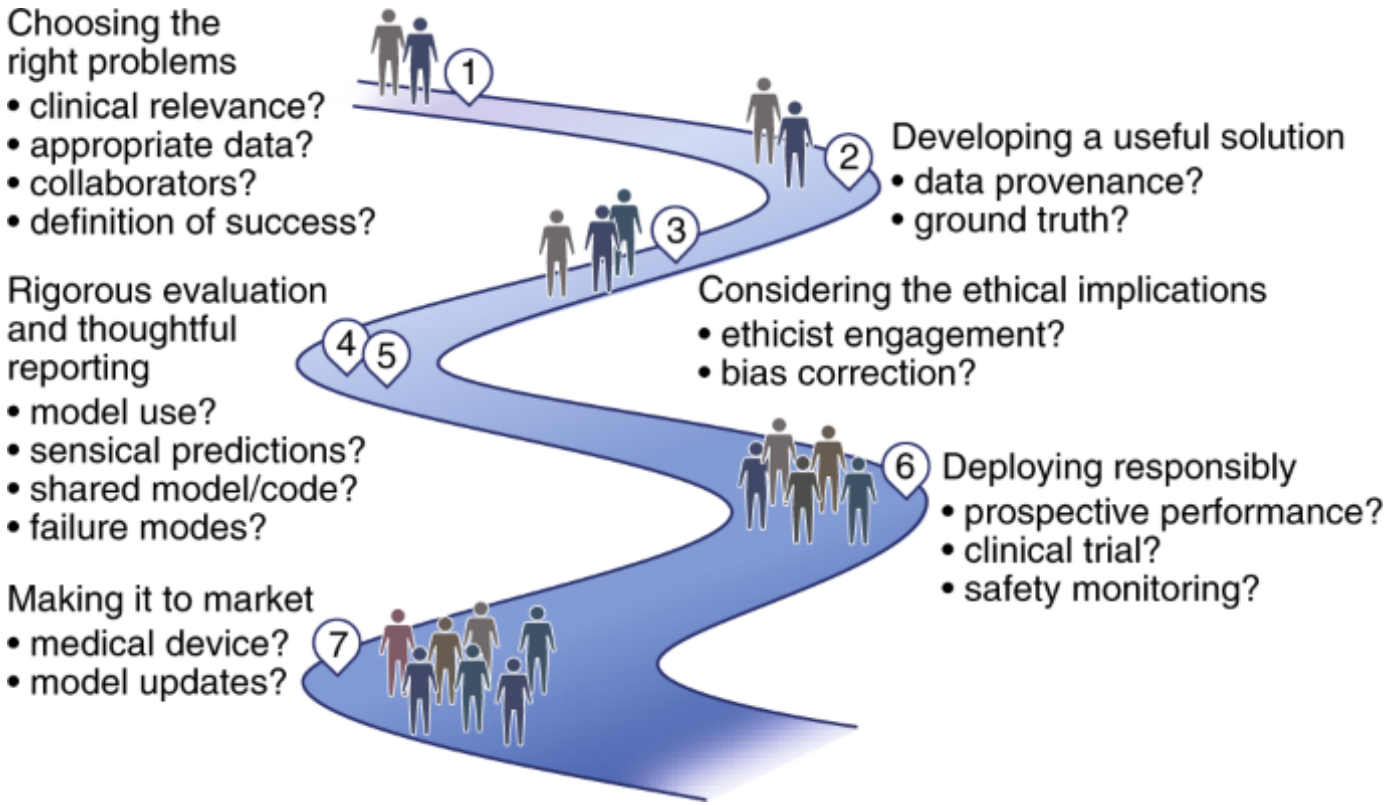
nature › nature medicine › perspectives › article

Perspective | Published: 19 August 2019

Do no harm: a roadmap for responsible machine learning for health care

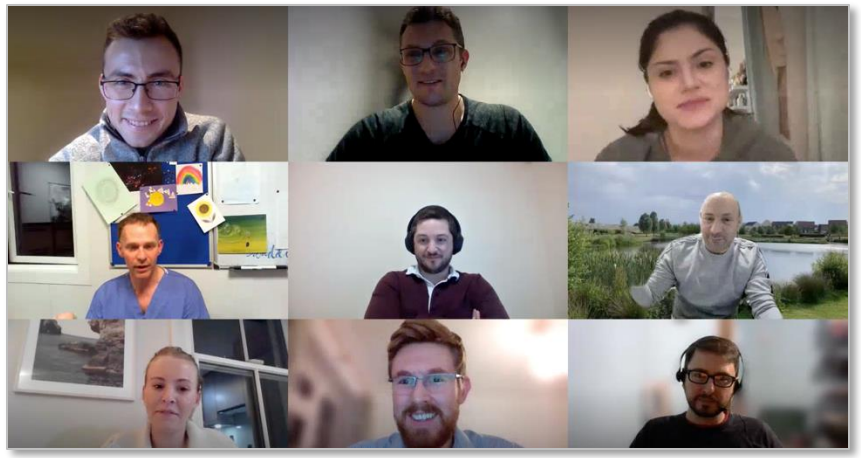
Jenna Wiens ✉, Suchi Saria, Mark Sendak, Marzyeh Ghassemi, Vincent X. Liu, Finale Doshi-Velez, Kenneth Jung, Katherine Heller, David Kale, Mohammed Saeed, Pilar N. Ossorio, Sonoo Thadaney-Israni & Anna Goldenberg ✉

Nature Medicine **25**, 1337–1340 (2019) | Cite this article



Wiens, J., Saria, S., Sendak, M. *et al.* Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* **25**, 1337–1340 (2019). https://doi.org/10.1038/s41591-019-0548-6

Patient and end user engagement is essential to understand workflows and problems.



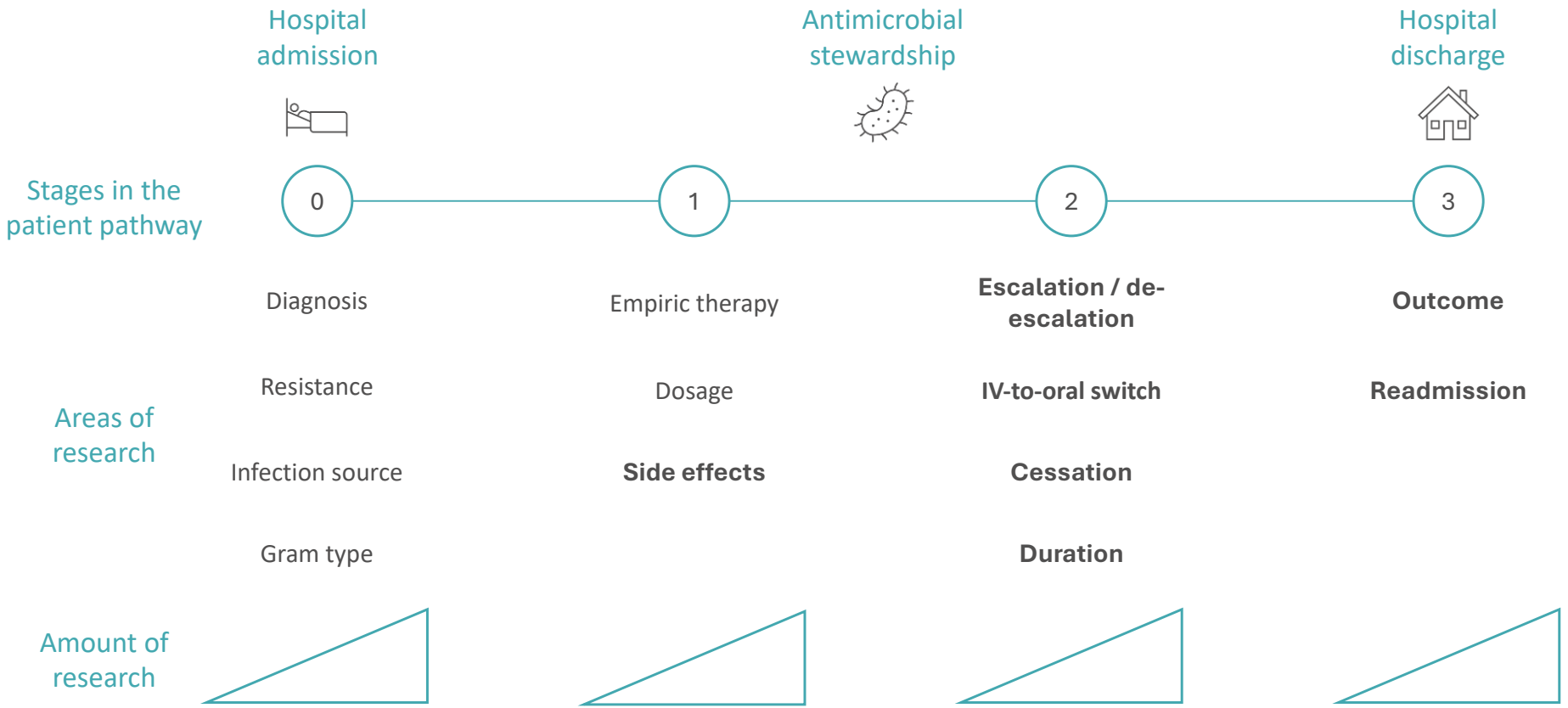
Bacterial antimicrobial resistance is a growing global threat.

5 deaths every hour in the UK

Costs the NHS **£230m** a year

>**30%** of antibiotic prescriptions are inappropriate

Antimicrobial stewardship aims to optimise antibiotic decision making.

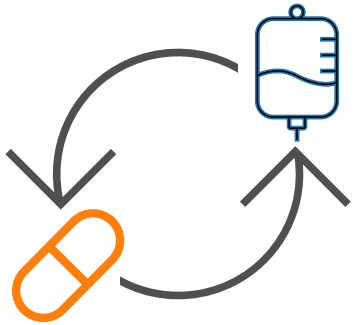


Antimicrobial stewardship

A coordinated effort and set of practices aimed at **optimising antimicrobial use** and **prolonging their therapeutic life**, to improve infection patient **outcomes** while minimizing the development of **antimicrobial resistance**

Oral antibiotic 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**

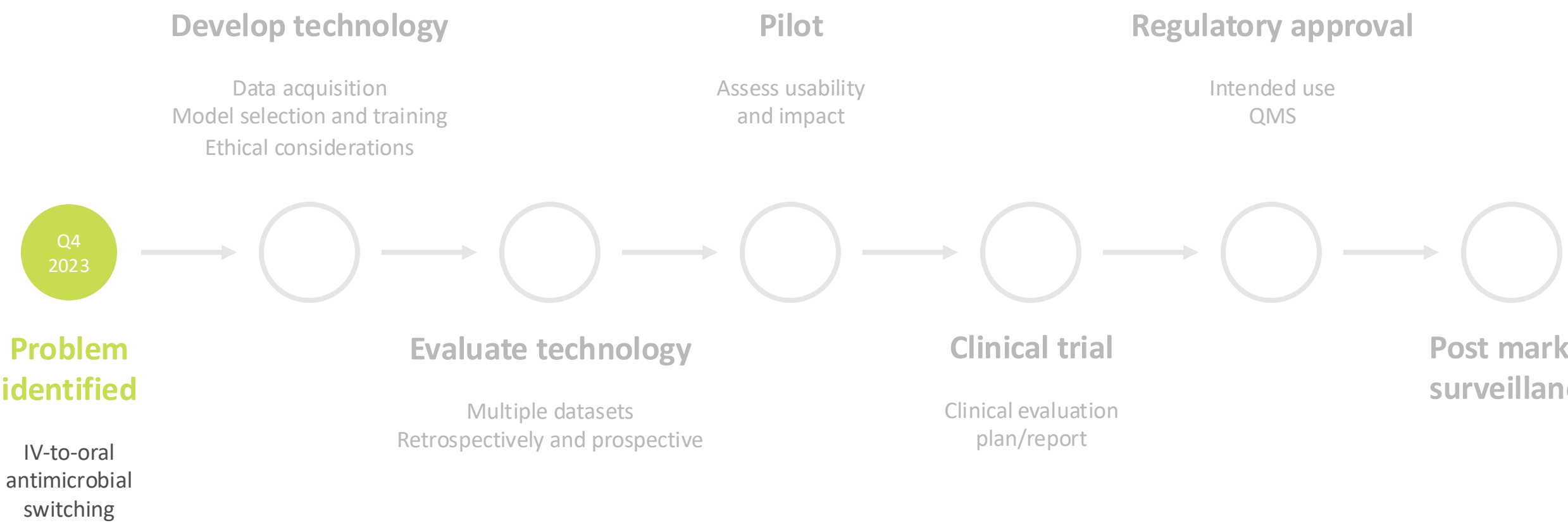
Oral therapy are often **non-inferior** to IV with **fewer side effects**, decreased **nursing workload**, lower **costs**, reduced **climate impact** and improved **patient comfort**

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

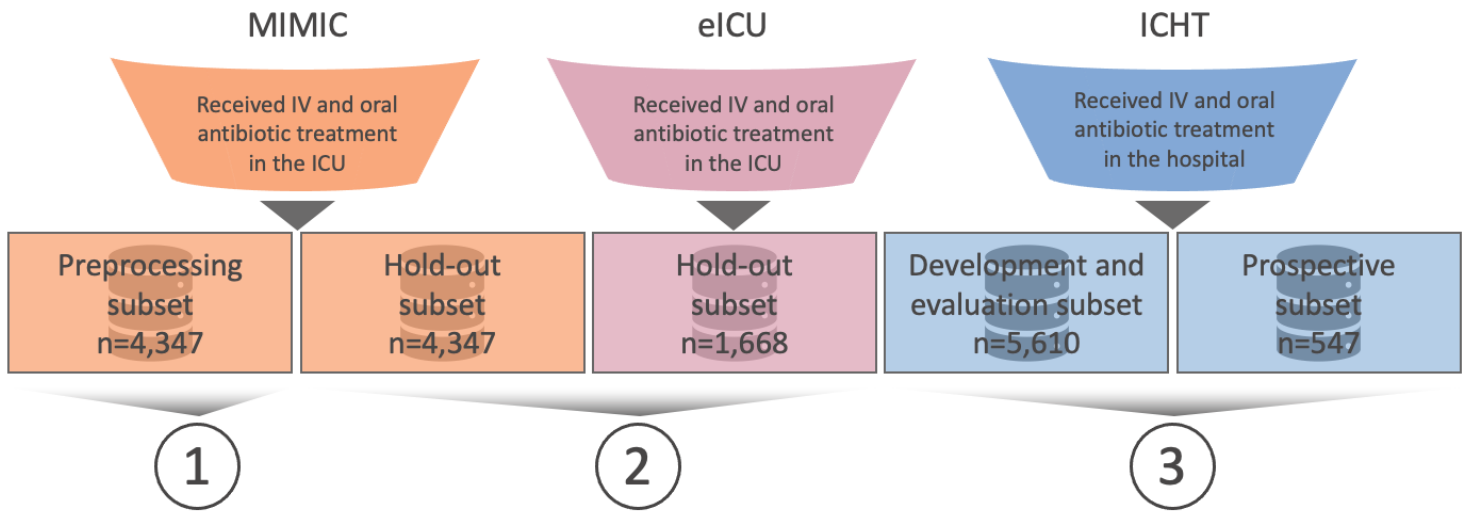
Hypothesis

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**

The roadmap for AI clinical decision support systems.



Machine learning models were trained to predict a patients route of administration in 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

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

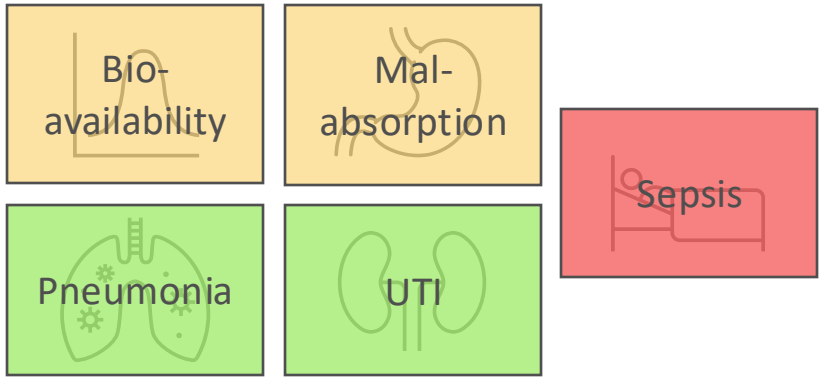
eICU

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	Results	Prospective dataset	Prospective PPS
AUROC	0.79 (SD 0.01)	0.77	0.68
FPR	0.21 (SD 0.03)	0.20	0.28

SUBGROUPS



nature communications

The roadmap for AI clinical decision support systems.

Problem identified

IV-to-oral antimicrobial switching

Q3 2022

Developed technology

Published in leading academic journals and conferences

nature communications · nature machine intelligence · AAAI Association for the Advancement of Artificial Intelligence · ICLR International Conference On Learning Representations

Q1 2023

Evaluated technology

MIMIC, eICU and ICHT datasets

Q3 2023

Pilot

Assess usability and impact

Clinical trial

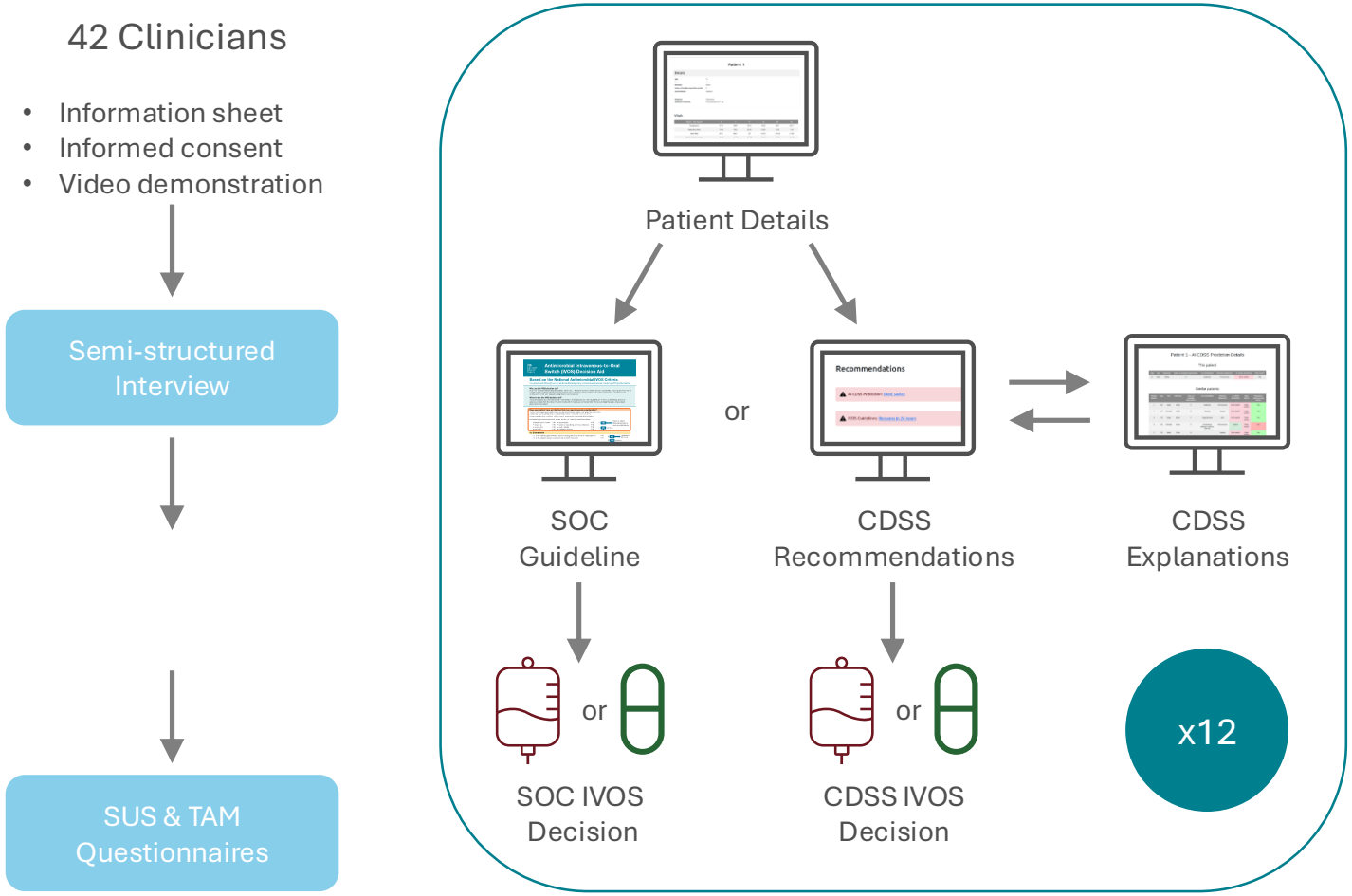
Clinical evaluation plan/report

Regulatory approval

Intended use QMS

Post market surveillance

A clinician evaluation study was conducted with case vignettes, interviews and questionnaires.



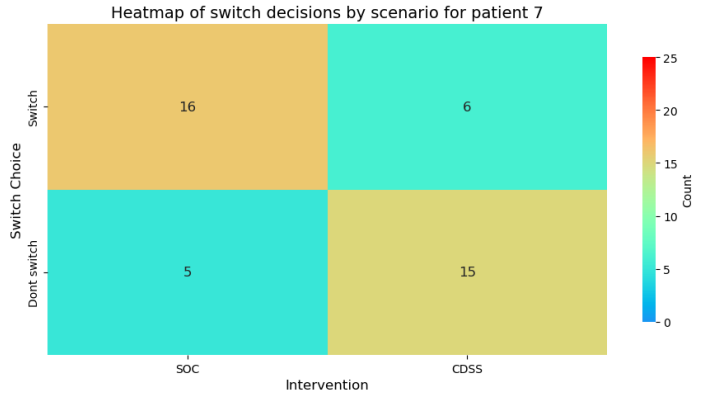
A greater impact was observed when the AI-IVOS CDSS recommended don't switch.

NO DIFFERENCES

11/12 cases

SIGNIFICANT DIFFERENCES

Patient 7 and globally with a GLME model

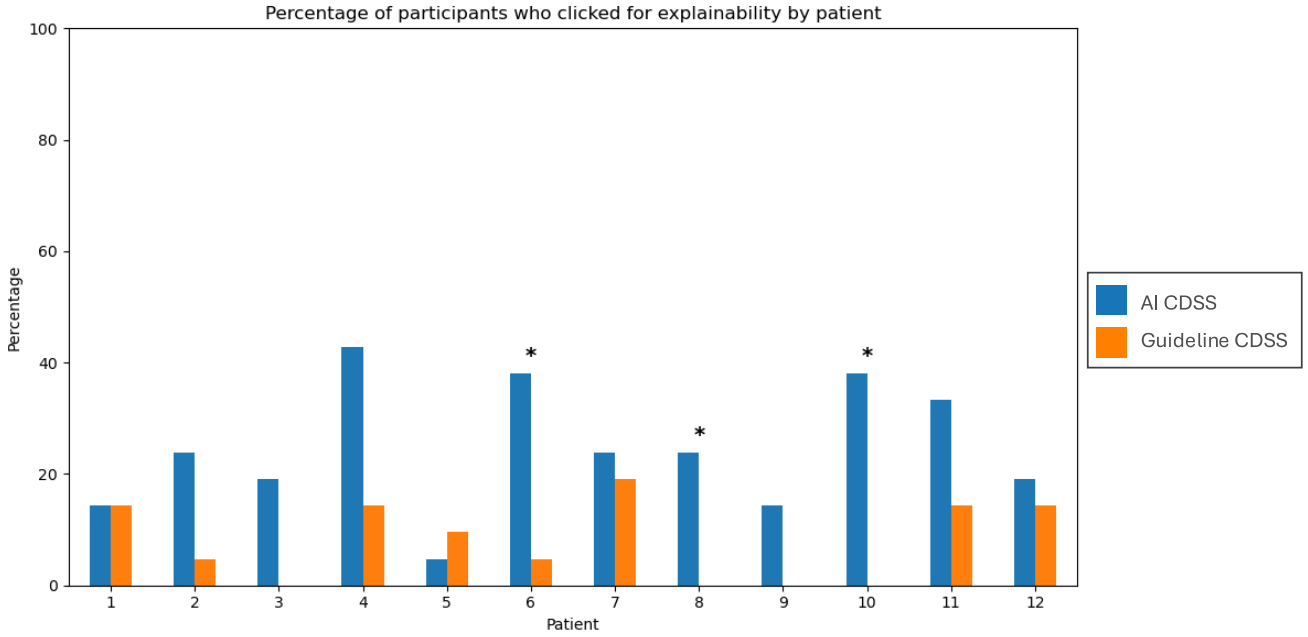


System Usability score: 72.32 / 100

Perceived usefulness: 3.59 / 5

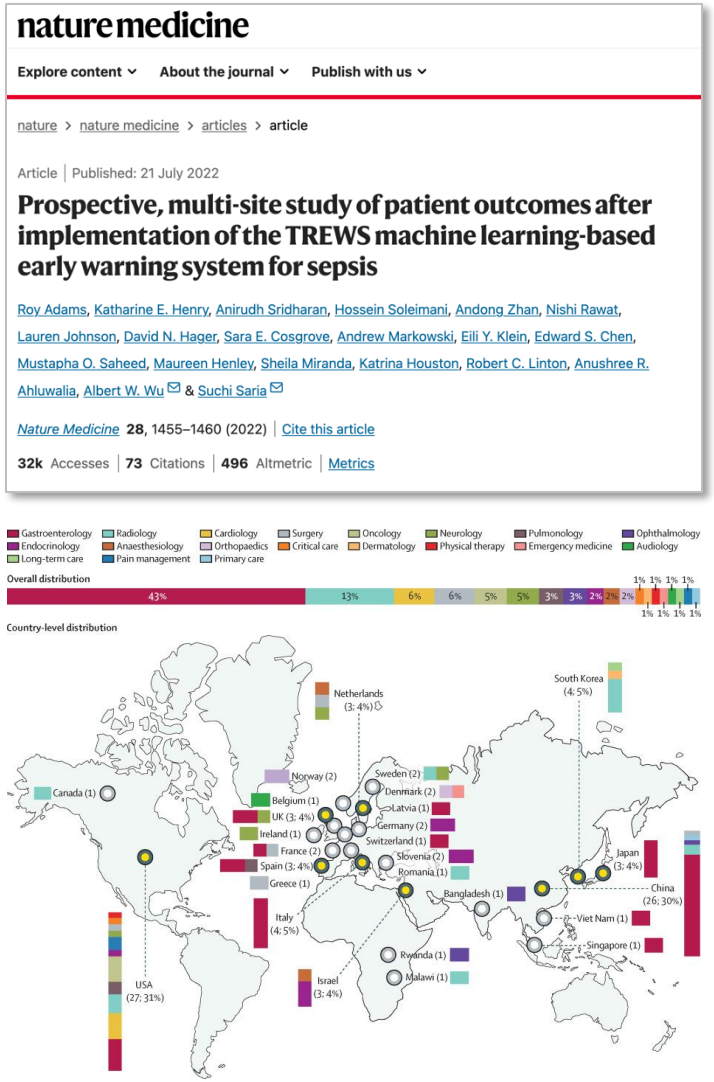
Perceived ease of use: 3.83 / 5

Self efficacy: 4.05 / 5

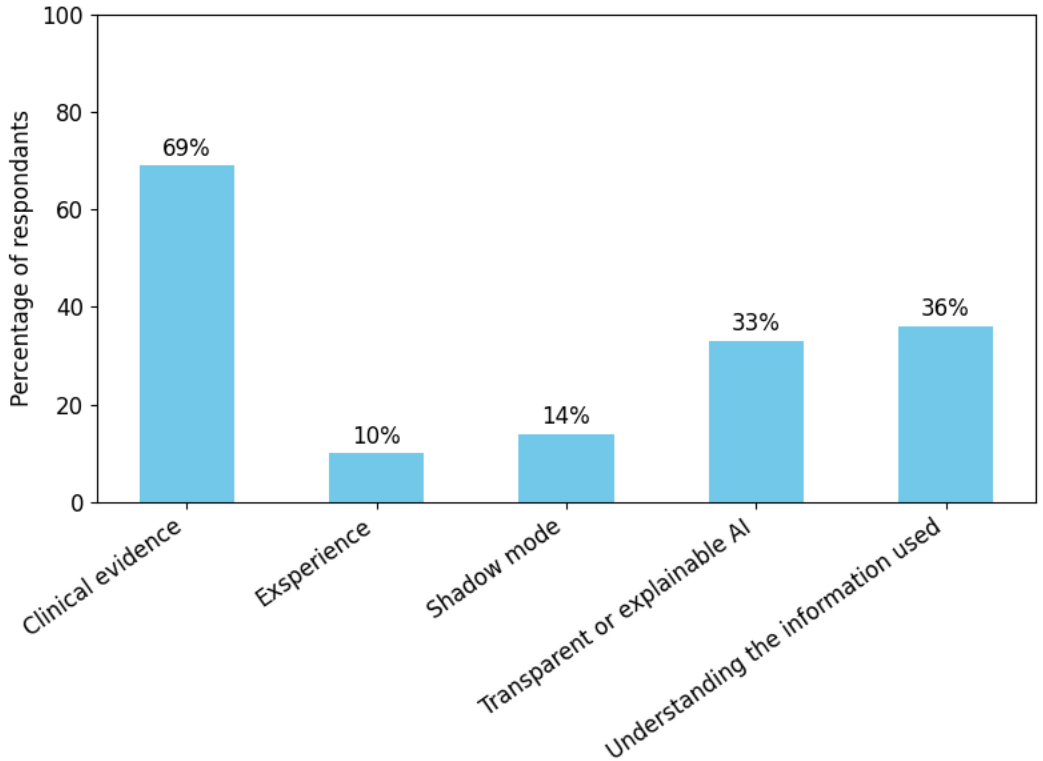


AI CDSS explanations were only viewed **9%** of the time, with those cases when the AI recommended "Don't switch" seeing explanations viewed more (e.g., patients 4, 6 and 10)

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



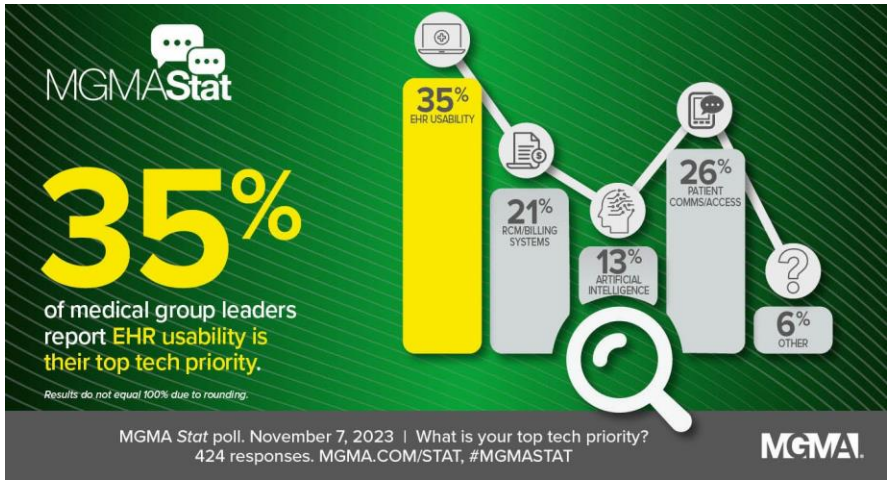
What factors determine the level of trust you have in an AI system?



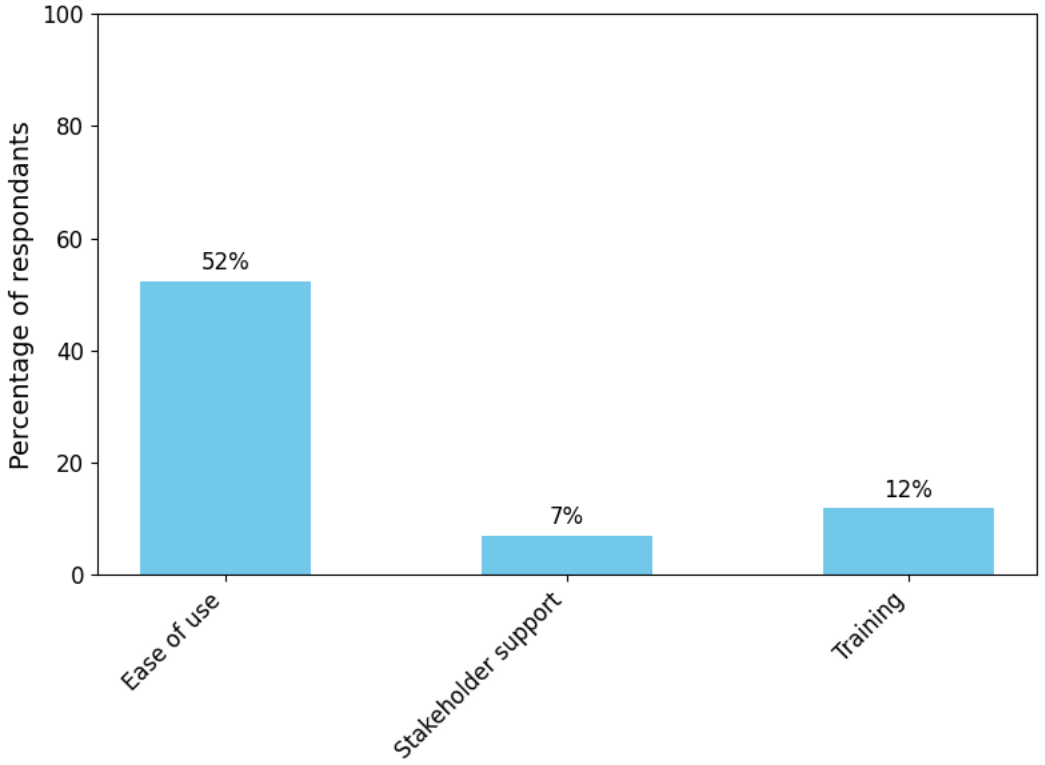
"Demonstration of utility and safety is absolutely critical."

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?

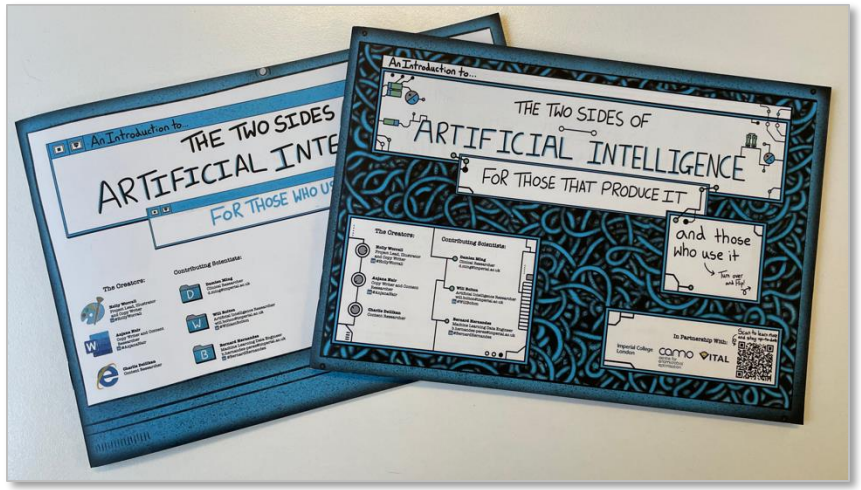


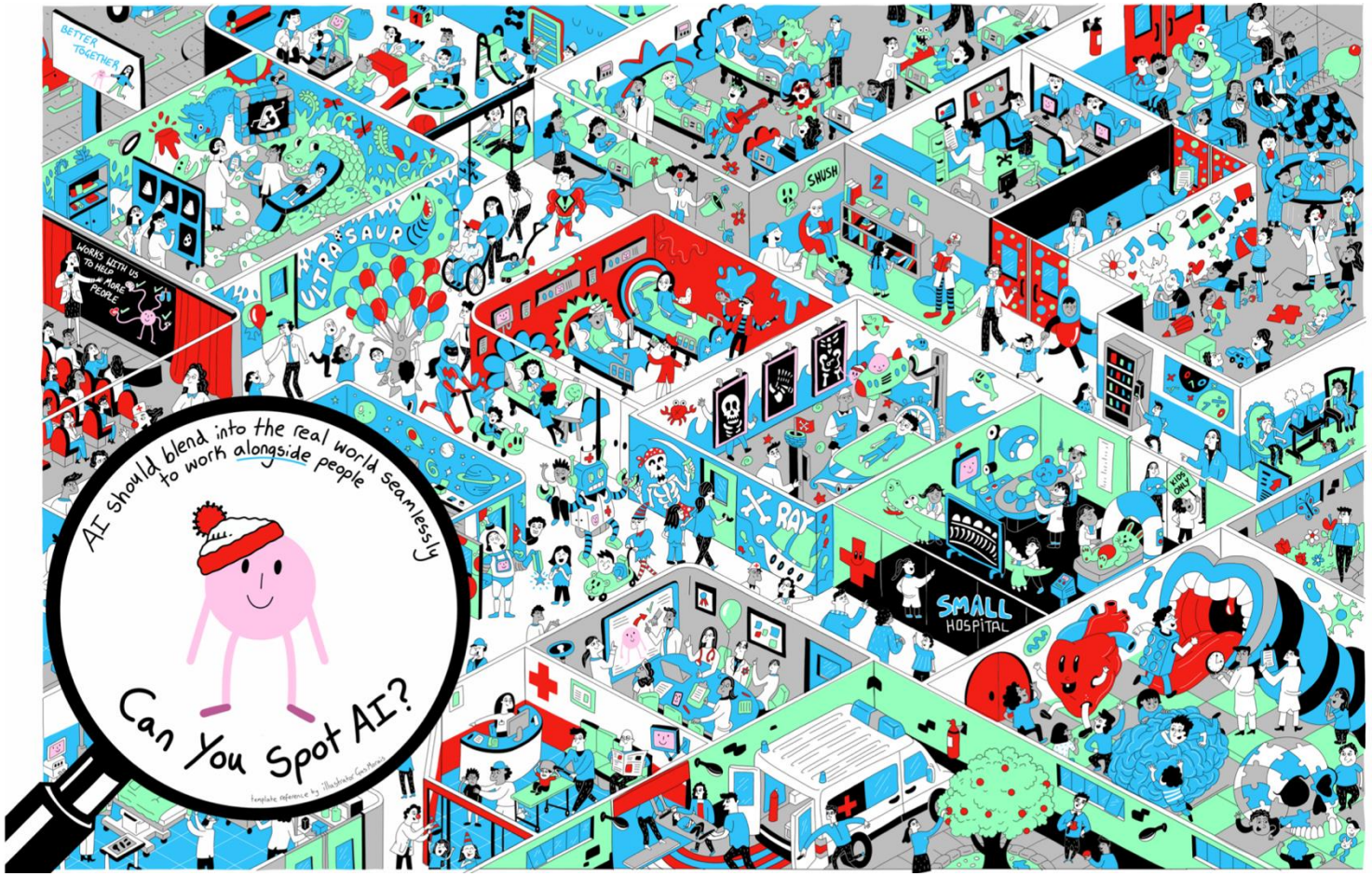
"I think the UI UX, and the whole design in terms of making it intuitive, is key."

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.

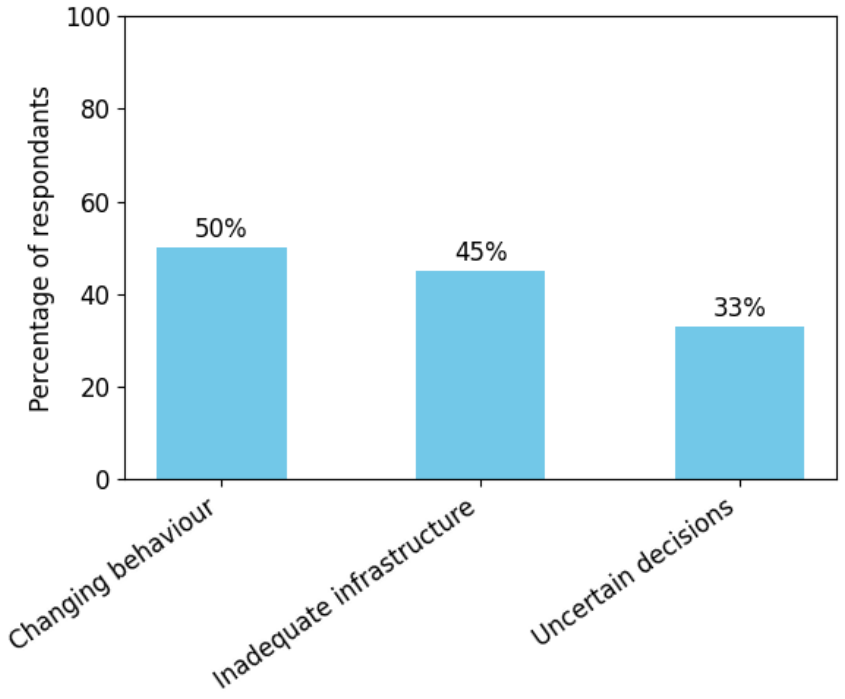
Education and training are important to ensure any new technology is adopted and used appropriately.



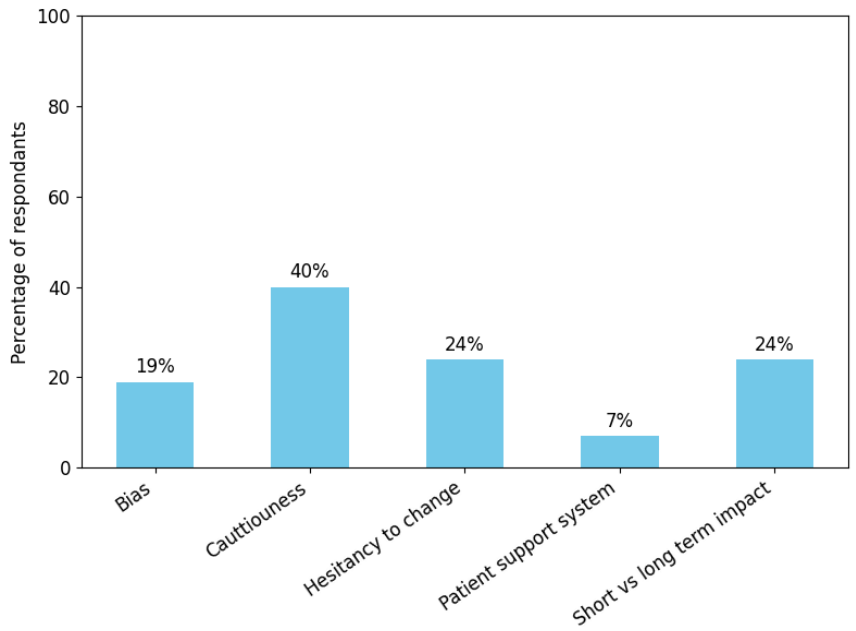


Infrastructure, behaviour, culture and uncertainty pose challenges for AI systems in healthcare, particularly those focusing on AMR.

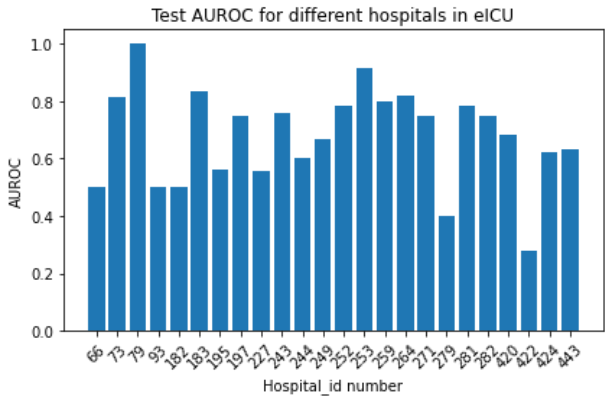
What is the biggest bottleneck in translating healthcare-focused AI research?



What factors influence antibiotic stewardship decisions?

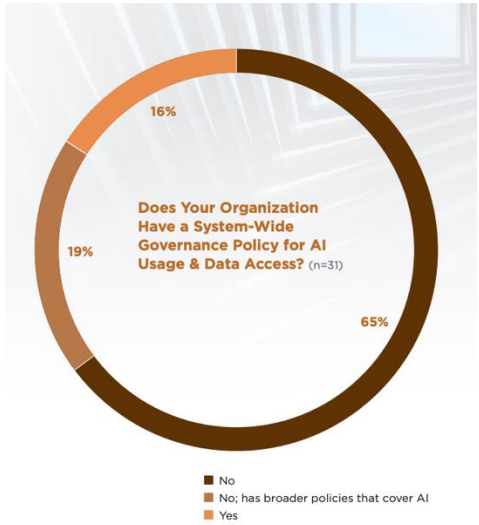


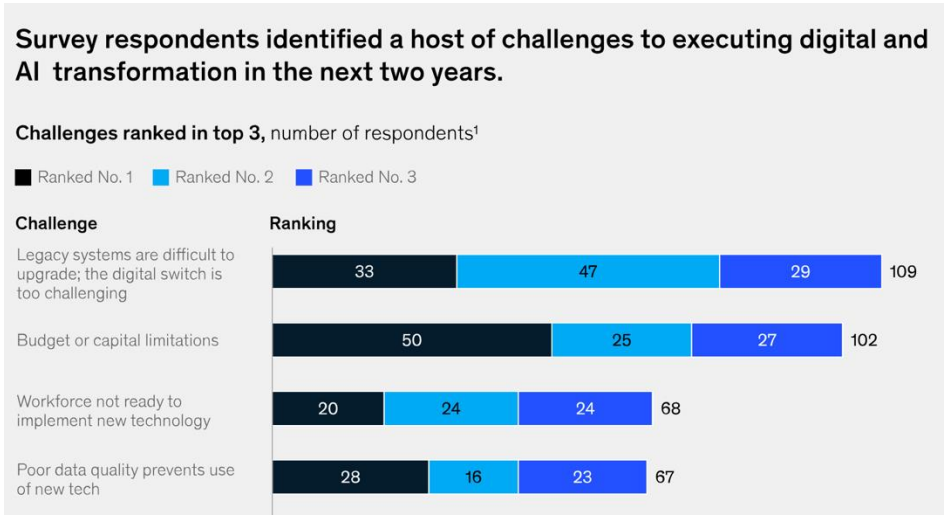
HUMAN BEHAVIOUR IS HETEROGENEOUS



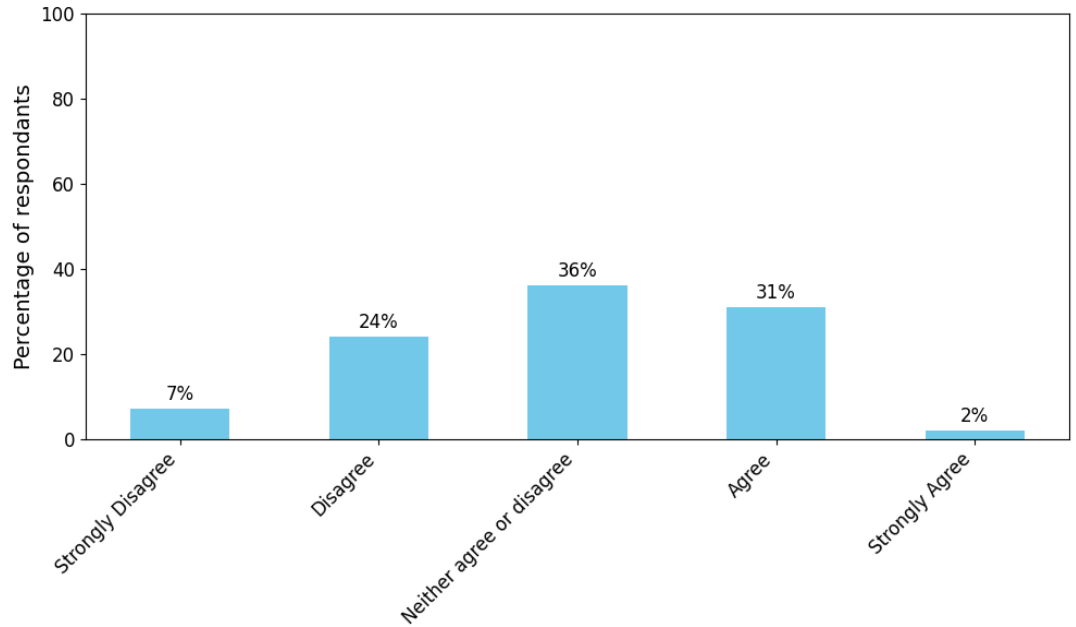
DATA QUALITY AND MISSINGNESS

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

The roadmap for AI clinical decision support systems.

Problem identified

IV-to-oral antimicrobial switching

Q3 2022

Developed technology

Published in leading academic journals and conferences

nature communications · nature machine intelligence · AAAI Association for the Advancement of Artificial Intelligence · ICLR International Conference On Learning Representations

Q1 2023

Evaluated technology

MIMIC, eICU and ICHT datasets

Q3 2023

Piloted

In simulated environment

conceptionx PhD DEEP TECH STARTUPS · AI SuperConnector

Q3 2024

Clinical trial

Clinical evaluation plan/report

Regulatory approval

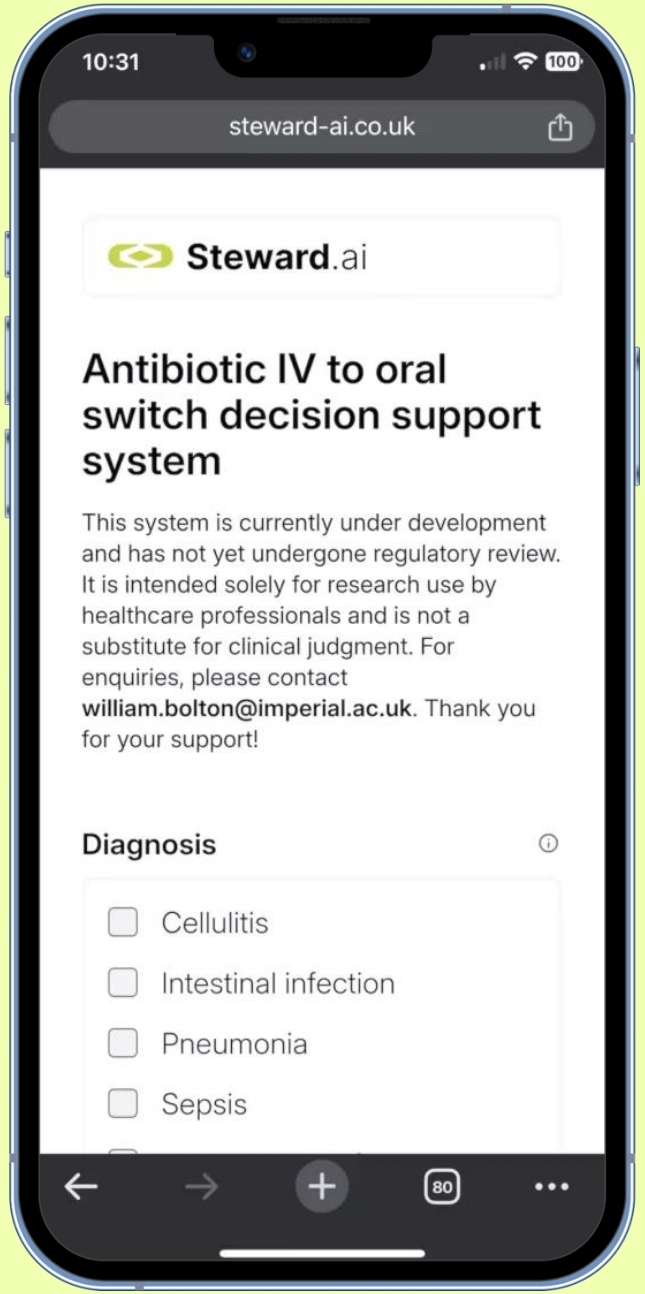
Intended use
QMS

Post market surveillance

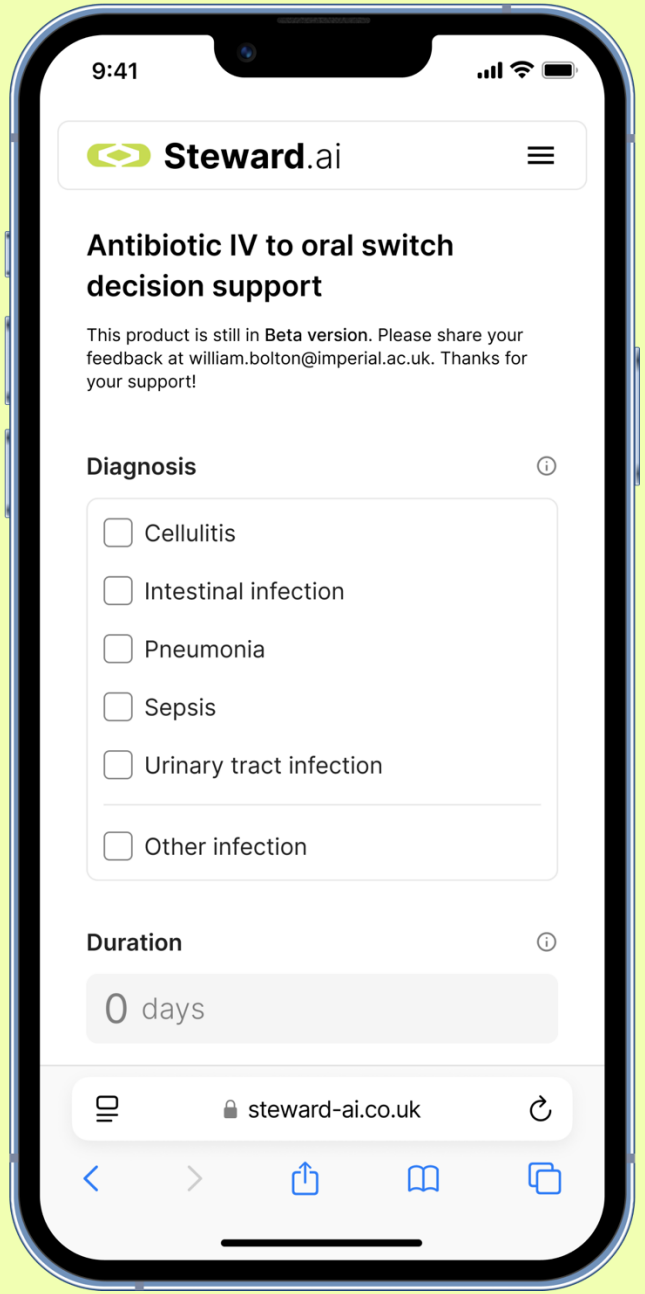
Steward.ai

Empowering Clinicians for
Precision Antibiotic Therapy

Steward.ai

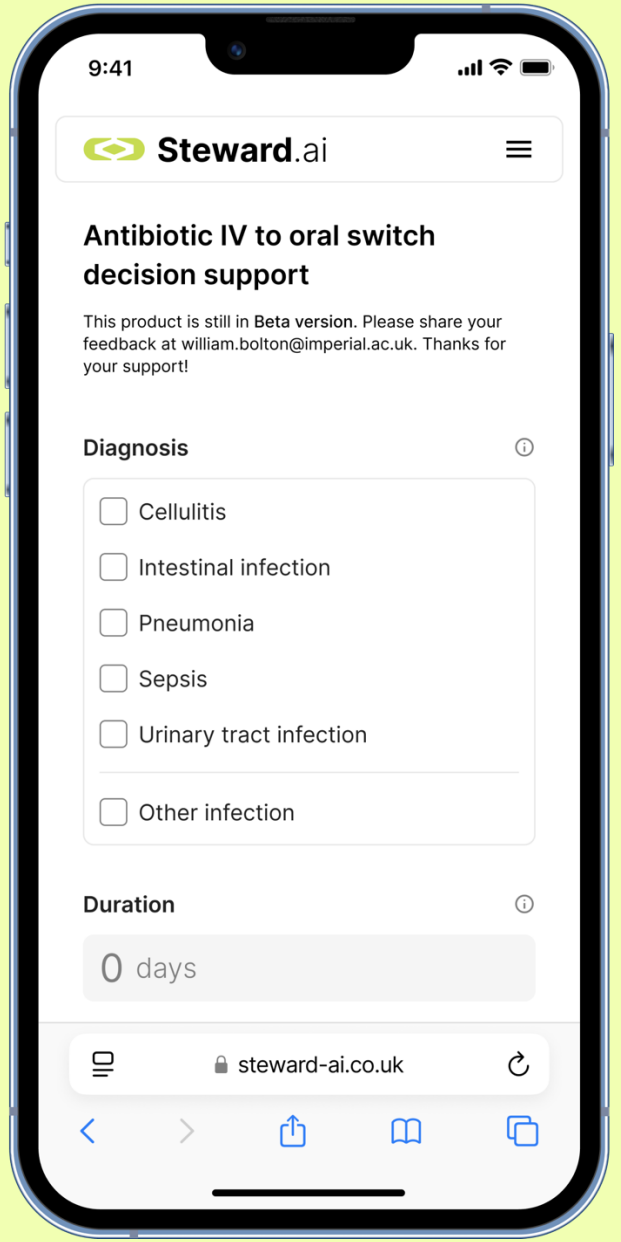

10:31
steward-ai.co.uk
Steward.ai
Antibiotic IV to oral switch decision support system
This system is currently under development and has not yet undergone regulatory review. It is intended solely for research use by healthcare professionals and is not a substitute for clinical judgment. For enquiries, please contact william.bolton@imperial.ac.uk. Thank you for your support!
Diagnosis
Cellulitis
Intestinal infection
Pneumonia
Sepsis

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 predictions
- Simple and **easy to use**
- Explainable, **safeguarded AI**
- Robust **clinical evaluation**



Steward.ai sees improved prospective evaluation performance with a reduced false positive rate.

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

Professor Mark Gilchrist

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

68 days of antibiotic treatment

Metric	Prior 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	-	0.81	-
Precision	-	0.85	-
Recall	-	0.58	-
F1 score	-	0.69	-
Weighted Partial Credit Accuracy	-	-	0.74
Weighted Cohen's Kappa	-	-	0.42

Steward.ai sees improved prospective evaluation performance with a reduced false positive rate.

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

Professor Mark Gilchrist

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 results	Steward.ai results
Before	37.5%	**8%**
Same day	33.5%	**58%**
After	**30%**	33%

The roadmap for AI clinical decision support systems.

Problem identified

Spoken to 50+ clinicians

Q3 2022

Developed technology

Web application for accessible decision support

AI SuperConnector

Q4 2024

Evaluated technology

ICHT and prospective point prevenance survey

Q1 2025

Pilot

In real clinical environment

NHS Imperial College Healthcare NHS Trust

Professor Mark Gilchrist

Q2 2025

Clinical trial

Define trial design, population and outcomes

Establish partnerships and licenses with other hospitals and MedTech businesses for commercialisation

Q1 2026

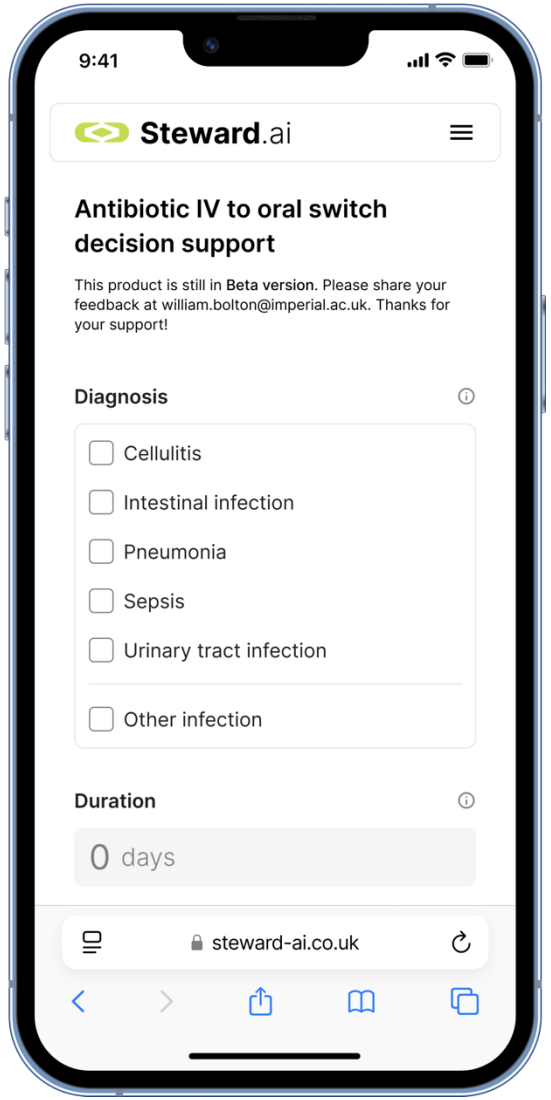
Regulatory approval

Intended use
QMS

Post market surveillance

Key takeaways when navigating clinical translation.

- Immerse yourself in the **clinical problem**
- Set up **interdisciplinary collaborations**
- **Plan ahead** (e.g., TRIPOD-AI or FUTURE-AI frameworks)
- Focus on **clinical evidence** and **usability**
- **Iterate** quickly and **learn** continually
- Be **resilient** it is a long journey



Thank you for the support.

Dr Tim Rawson

Professor Pantelis Georgiou

Professor Alison Holmes

Professor Mark Gilchrist

Richard Wilson

Dr David Antcliffe

Dr Bernard Hernandez Perez

Cosmin Badea

Britta Ross

Imperial College London

NHS Imperial College Healthcare NHS Trust

Health

camo centre for antimicrobial optimisation

AI SuperConnector

AI Decision Support for Antimicrobial Stewardship: Navigating Clinical Translation

Dr William Bolton

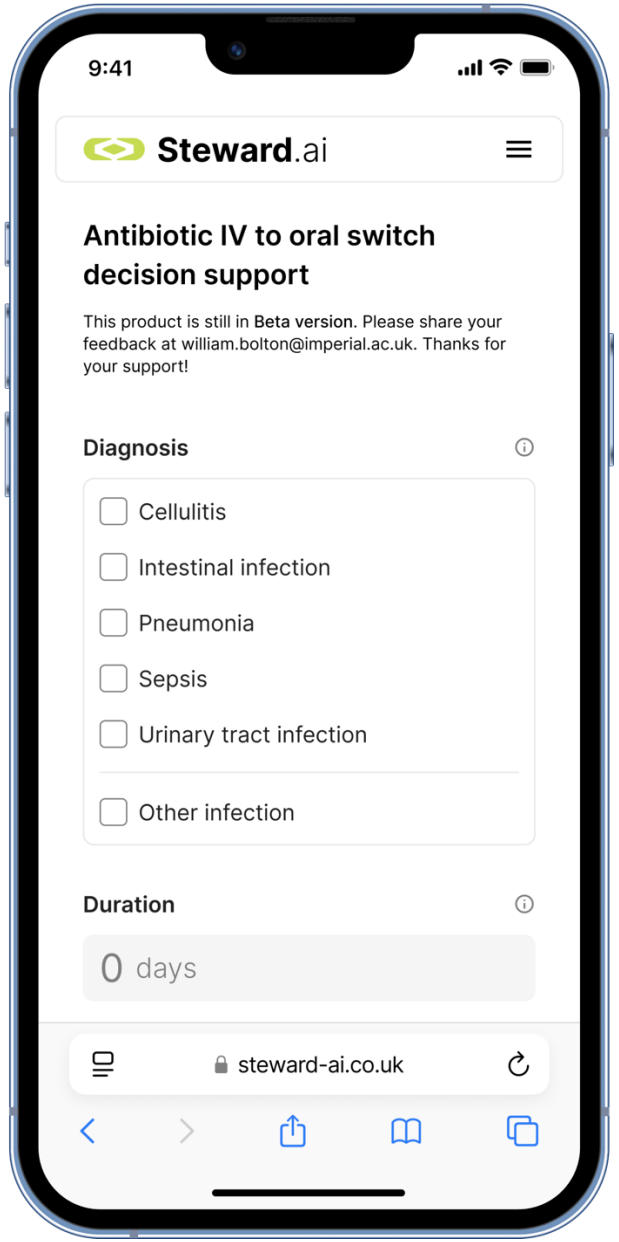
MedTech Links

24th March 2025

Personal Website

LinkedIn

Steward.ai



Artificial intelligence can support optimised antibiotic decision making.

STAGES OF ANTIBIOTIC DECISION MAKING

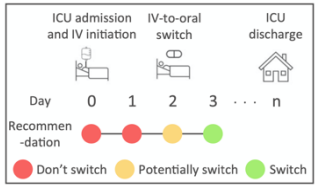
Hospital admission

Antimicrobial stewardship

Hospital discharge

0 — 1 — 2 — 3

IV-to-oral switch



Antibiotic readmission

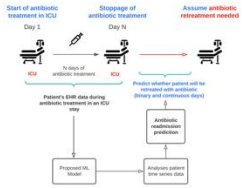


Figure 1.2: Proposed ML-based decision support model

Side effects

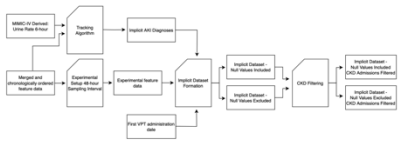
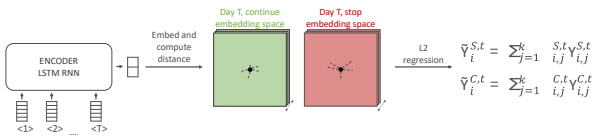


Figure 13: Implicit dataset formation workflow.

Antibiotic cessation



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

AI clinical decision support systems are often regulated as a Class II software as a medical device in the UK.

The Journey

Pre-Market

1. **Intended Use**
   The critical first step in the development of AI and health tech products. A clear intended use prioritises safety, effectiveness and gives clarity on how to position your SaMD for success.
   READ MORE →

2. **Risk Classification**
   The risk level of a medical device or AI product determines the required clinical evidence and regulatory oversight. Read on to learn more!
   READ MORE →

3. **Notified / Approved Body Engagement**
   To launch a product in the UK or EU, an independent notified body or approved body must review and comply with European legislation, granting UKCA marks.
   READ MORE →

4. **Quality Management Systems (ISO 13485)**
   Elevate compliance in medical device manufacturing with our QMS. We cover design, supply, risk management, and CAPAs for a solid regulatory strategy. Read more!
   READ MORE →

5. **Medical Device File Design**
   Your MDF provides evidence to demonstrate compliance of the device to all the applicable regulations. Its structure and design is key.
   READ MORE →

6. **MDSAP**
   The Medical Device Single Audit Programme (MDSAP) streamlines quality management systems by enabling compliance proof through a single audit for five markets.
   READ MORE →

7. **Clinical Evaluation Plan**
   The Clinical Evaluation Plan (CEP) is a vital tool in product development, guiding device clinical evaluation through the Valid Clinical Association. Read more!
   READ MORE →

8. **QMS Deployment & Training**
   Hardian is an AI-based SaMD company that offers training and guidance to help clients meet international standards for quality management system implementation.
   READ MORE →

9. **Software Verification & Validation**
   The process of the software development lifecycle is essential to ensure all requirements are met before testing the product in the real world. Read more!
   READ MORE →

10. **Clinical Evaluation Report**
    The Clinical Evaluation Report (CER) is a crucial document in clinical evaluation, containing development activities and clinical evidence for device marketing.
    READ MORE →

11. **Responsible Person (PRRC and UKRP)**
    Depending on your jurisdiction of deployment, you may need to appoint responsible persons across these jurisdictions.
    READ MORE →

12. **Product Registration**
    Now the hard work is done, registering your product on the relevant regulatory databases is the final step to legally place your product on the market.
    READ MORE →

Medicines & Healthcare products Regulatory Agency

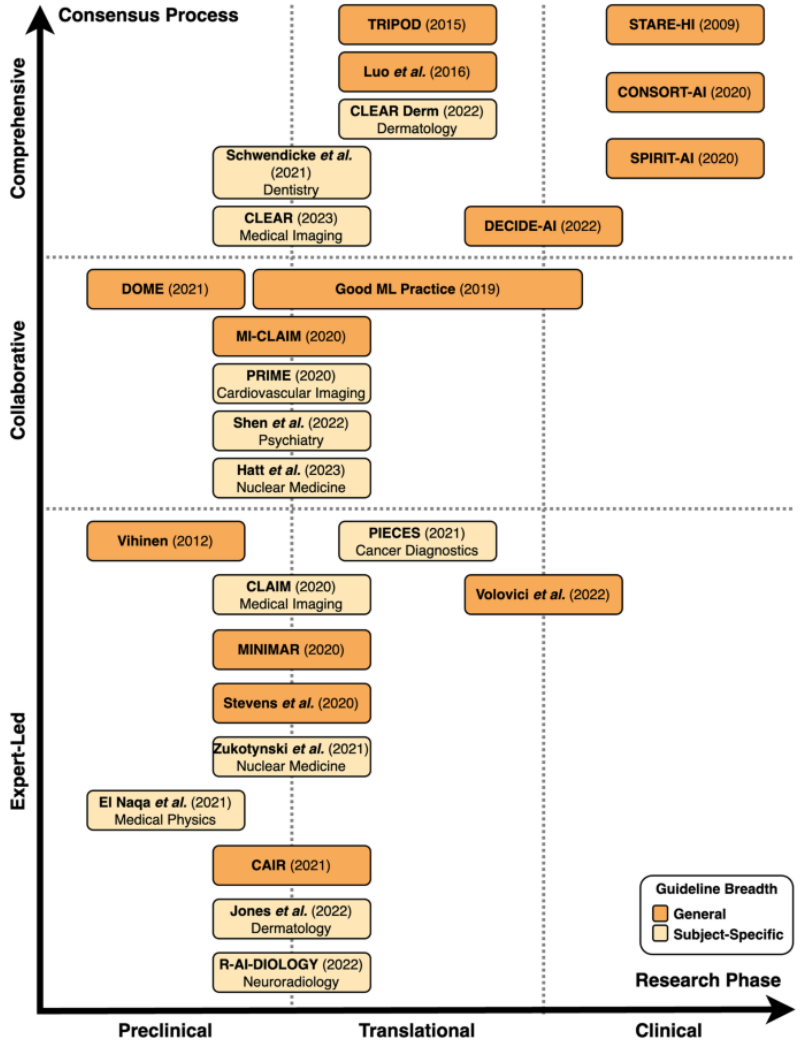
Guidance

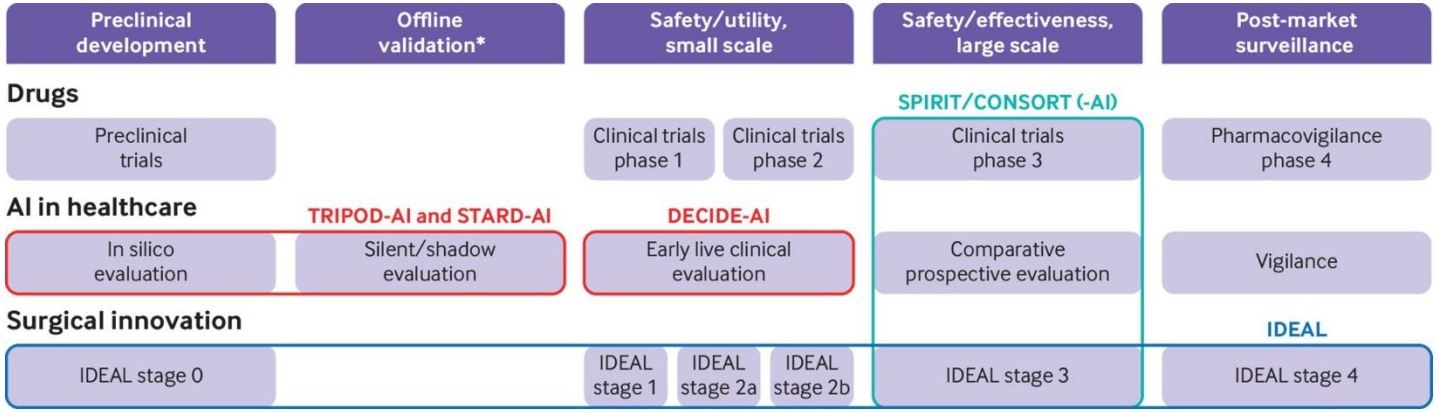
Software and artificial intelligence (AI) as a medical device

Updated 13 June 2024

https://www.england.nhs.uk/long-read/artificial-intelligence-ai-and-machine-learning/
https://www.hardianhealth.com/regulatory
https://www.gov.uk/government/publications/software-and-artificial-intelligence-ai-as-a-medical-device/software-and-artificial-intelligence-ai-as-a-medical-device

Many guidelines exist for reporting AI in medicine.





Box 2: Noteworthy changes and additions to TRIPOD 2015

- New checklist of reporting recommendations to cover prediction model studies using any regression or machine learning method (eg, random forests, deep learning), and harmonise nomenclature between regression and machine learning communities
- New TRIPOD+AI checklist supersedes the TRIPOD 2015 checklist, which should no longer be used
- Particular emphasis on fairness (box 1) to raise awareness and ensure that reports mention whether specific methods were used to deal with fairness. Aspects of fairness are embedded throughout the checklist
- Inclusion of TRIPOD+AI for Abstracts for guidance on reporting abstracts
- Modification of the model performance item recommending that authors evaluate model performance in key subgroups (eg, sociodemographic)
- Inclusion of a new item on patient and public involvement to raise awareness and prompt authors to provide details on any patient and public involvement during the design, conduct, reporting (and interpretation), and dissemination of the study
- Inclusion of an open science section with subitems on study protocols, registration, data sharing and code sharing

TRIPOD=Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis; AI=artificial intelligence.

Medicines & Healthcare products Regulatory Agency

Guidance

Good Machine Learning Practice for Medical Device Development: Guiding Principles

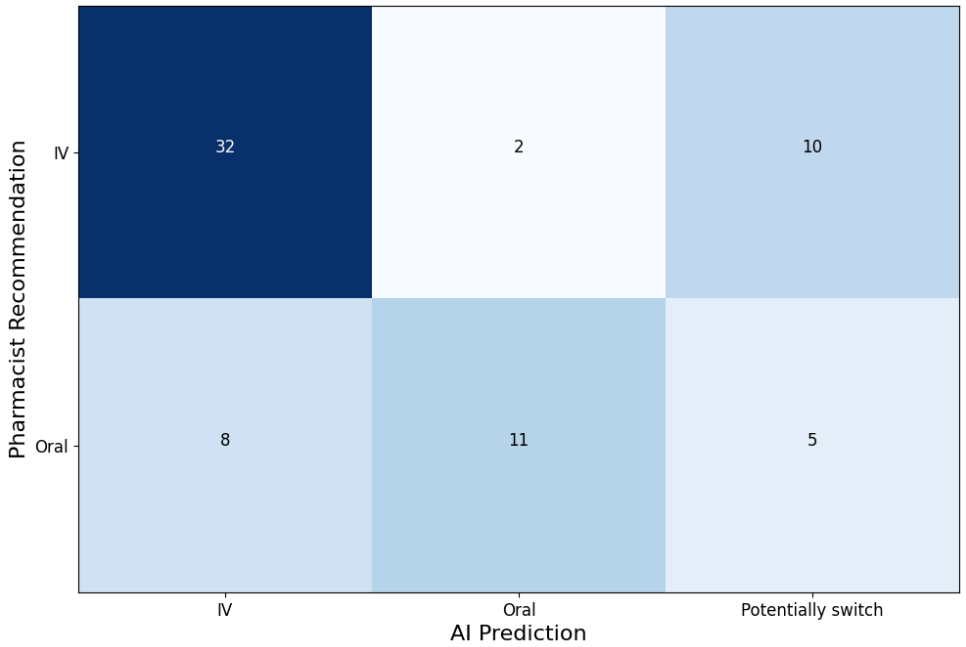
Published 27 October 2021

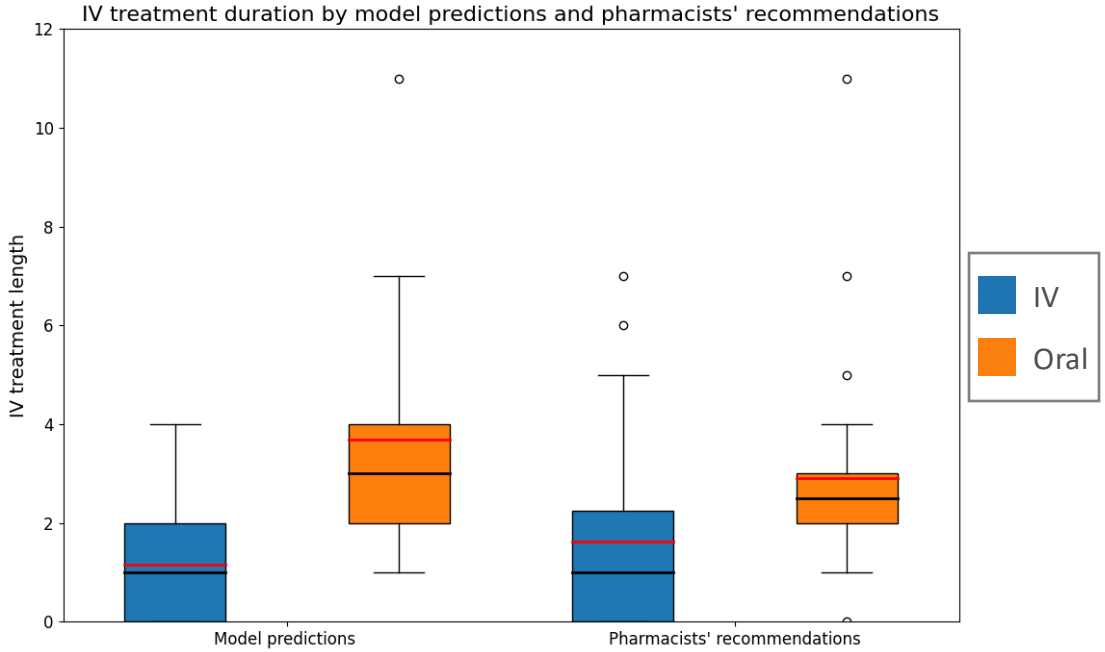
Kolbinger, F.R., Veldhuizen, G.P., Zhu, J. *et al.* Reporting guidelines in medical artificial intelligence: a systematic review and meta-analysis. *Commun Med* **4**, 71 (2024). https://doi.org/10.1038/s43856-024-00492-0

Vasey B, Nagendran M, Campbell B, Clifton D A, Collins G S, Denaxas S et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI BMJ 2022; 377 :e070904 doi:10.1136/bmj-2022-070904

Collins G S, Moons K G M, Dhiman P, Riley R D, Beam A L, Van Calster B et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods BMJ 2024; 385 :e078378 doi:10.1136/bmj-2023-078378

Steward.ai sees improved prospective evaluation performance with a reduced false positive rate.





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