The impact of artificial intelligence-driven decision support on uncertain antimicrobial prescribing: a randomised, multimethod study

William J Bolton, Richard Wilson, Mark Gilchrist, Pantelis Georgiou, Alison Holmes, Timothy M Rawson

Summary

Background Challenges exist when translating artificial intelligence (AI)-driven clinical decision support systems (CDSSs) from research into health-care settings, particularly in infectious diseases, an area in which behaviour, culture, uncertainty, and frequent absence of a ground truth enhance the complexity of medical decision making. We aimed to evaluate clinicians' perceptions of an AI CDSS for intravenous-to-oral antibiotic switching and how the system influences their decision making.

Methods This randomised, multimethod study enrolled health-care professionals in the UK who were regularly involved in antibiotic prescribing. Participants were recruited through personal networks and the general email list of the British Infection Association. The first part of the study involved a semistructured interview about participants' experience of antibiotic prescribing and their perception of AI. The second part used a custom web app to run a clinical vignette experiment: each of the 12 case vignettes consisted of a patient currently receiving intravenous antibiotics, and participants were asked to decide whether or not the patient was suitable for switching to oral antibiotics. Participants were assigned to receive either standard of care (SOC) information, or SOC alongside our previously developed AI-driven CDSS and its explanations, for each vignette across two groups. We assessed differences in participant choices according to the intervention they were assigned, both for each vignette and overall; evaluated the aggregate effect of the CDSS across all switching decisions; and characterised the decision diversity across participants. In the third part of the study, participants completed the system usability scale (SUS) and technology acceptance model (TAM) questionnaires to enable their opinions of the AI CDSS to be assessed.

Findings 59 clinicians were directly contacted or responded to recruitment emails, 42 of whom from 23 hospitals in the UK completed the study between April 23, 2024, and Aug 16, 2024. The median age of participants was 39 years (IQR 37–47), 19 (45%) were female and 23 (55%) were male, 26 (62%) were consultants and 16 (38%) were training-grade doctors, and 14 (33%) specialised in infectious diseases. Interviews revealed mixed individualisation of prescribing and uneven use of technology, alongside enthusiasm for AI, which was conditional on evidence and usability but constrained by behavioural inertia and infrastructure limitations. Case vignette completion times and many decisions were equivalent between SOC and CDSS interventions, with clinicians able to identify and ignore incorrect advice. When a statistical difference was observed, the CDSS influenced participants towards not switching (χ^2 7·73, p=0·0054; logistic regression odds ratio 0·13 [95% CI 0·03–0·50]; p=0·0031). AI explanations were used only 9% of the time when available. Our software and AI CDSS obtained a good SUS score of 72·3 out of 100 (SD 8·79) and, for the TAM questionnaire, scores of 3·6 out of 5 (0·31) for perceived usefulness, 3·8 out of 5 (0·20) for perceived ease of use, and 4·1 out of 5 (0·05) for self-efficacy.

Interpretation This AI CDSS was positively received and has the potential to support antimicrobial prescribing, with the greatest influence on clinicians when it recommended not switching from intravenous to oral treatment. Further prospective research is needed to gather safety and benefit data and to understand behavioural changes as AI CDSSs enter clinical practice. Our research suggests that AI explanations are likely to have a minor role at the point of care, and that AI CDSS adoption and utilisation depends on systems being easy to use and trusted, primarily through clinical evidence.

Funding The UK Research and Innovation Centre for Doctoral Training in AI for Healthcare, and the National Institute for Health and Care Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College London.

Lancet Digit Health 2025; 7: 100912

Published Online December 9, 2025
https://doi.org/10.1016/j.landig.2025.100912

Department of Computing (W J Bolton PhD), **AI4Health Centre for Doctoral Training** (W J Bolton), **National Institute for Health Research, Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance** (W J Bolton, R Wilson MPharm, Prof M Gilchrist MSc, Prof P Georgiou PhD, Prof A Holmes MD, T M Rawson PhD), **Centre for Antimicrobial Optimisation** (W J Bolton, R Wilson, Prof M Gilchrist, Prof P Georgiou, Prof A Holmes, T M Rawson), **Centre for Bio-inspired Technology, Department of Electrical and Electronic Engineering** (Prof P Georgiou), **and Department of Infectious Diseases** (Prof A Holmes), **Imperial College London, London, UK; Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK** (R Wilson, Prof A Holmes); **Imperial College Healthcare NHS Trust, London, UK** (Prof M Gilchrist)

Correspondence to:
Dr William J Bolton, Centre for Antimicrobial Optimisation, Imperial College London, London W12 0NN, UK
william.bolton@imperial.ac.uk

Introduction

Artificial intelligence (AI)-based clinical decision support systems (CDSSs) have the potential to revolutionise health care.[1] However, few technologies have been truly translated, integrated, and adopted in clinical practice,[2–6] particularly those that tackle infectious disease and the growing problem of antimicrobial resistance (AMR) through improving antibiotic prescribing.[7,8] This speciality has numerous complexities compared with areas in which AI uptake has been more successful to date, such as radiology diagnosis.[9] For example, there is no clear gold-standard ground truth; decisions are often made in uncertain scenarios with little information; behavioural factors, such as risk appetite, have a key role in decision making; clinical practice is heterogeneous; and a balance is needed between the benefits and risks to an individual versus those of wider society.[10–13]

To cross the translation gap from research to real-world clinical use, robust evaluations and clinical evidence are necessary.[14–17] Numerous aspects of AI CDSSs must be considered, including who the end user is, what the intended use is, and how and when predictions are presented. Furthermore, the potential effects of the tool and any explanation it provides must be understood. In previous research, we developed an AI CDSS to predict when a patient could be suitable for switching from intravenous to oral antibiotics.[18] Our predictive models were trained on two datasets, Medical Information Mart for Intensive Care

Research in context

Evidence before this study

We conducted four searches on PubMed, with no language restrictions, covering the period from database inception to Aug 26, 2024. The first search was for the effect of an artificial intelligence (AI)-based clinical decision support system (CDSS) on antimicrobial stewardship, using the search terms "impact", "artificial intelligence", "decision support", and "antimicrobial stewardship". The second search was for the perception of an AI CDSS on antimicrobial stewardship, with the search terms: "perception", "artificial intelligence", "decision support", and "antimicrobial stewardship". The third and fourth searches were the same as the first and second, but with the search term "antimicrobial stewardship" replaced with "intravenous-to-oral switch". The first search returned six results. Three were literature reviews reporting how AI and machine learning can support antimicrobial prescribing, predict antimicrobial resistance, and be applied to infectious diseases more broadly. One study from South Korea reported how antimicrobial stewardship programmes (not using AI) can be implemented in resource-limited settings. Another study combined machine learning with a flow cytometry-assisted antimicrobial susceptibility test. The third study was published by our research group and reported the evaluation of a case-based reasoning algorithm with data from real clinical practice. This algorithm predicts an appropriate empirical antimicrobial agent and is not focused on the later stages of stewardship, such as intravenous-to-oral switching. The third search returned one result: a paper reporting a machine learning model for predicting the appropriate cyclosporine dose to give when switching from intravenous to oral administration in children who received a stem-cell transplant. The authors reported good performance in their hold-out test dataset, but noted that the effect of the model in a clinical context needed to be ascertained. Other studies have investigated the effect of AI-based CDSSs in different areas of medical specialty, such as radiology, pulmonology, and critical care. The second and fourth searches returned no results, highlighting a lack of research that robustly investigates users' perceptions of such technology, although we are aware of a few broader studies that highlighted clinicians' perspectives and opinions on AI CDSSs for antimicrobial prescribing. We are not aware of any studies using multimethod approaches to investigate both user perceptions and the potential effect of an AI CDSS on antimicrobial prescribing.

Added value of this study

To our knowledge, this study is the first to investigate how an AI CDSS can affect antimicrobial decision making. We provide direct evidence on how such CDSSs are used by clinicians when making decisions on antimicrobial prescribing. Contrary to previous studies, we use a three-method approach, gathering free-text perspectives with semistructured interviews, quantitative data with an interactive web app, and further information on usability and perceptions with two questionnaires. This robust approach enables us to triangulate our findings and show how a clinically focused CDSS can be comprehensively evaluated by its potential users. On a larger scale, our results contribute towards the evidence base for how AI CDSSs can be appropriately translated from the research setting into clinical practice.

Implications of all the available evidence

This study shows that our AI CDSSs can drive behaviour change in some antimicrobial prescribing situations, in particular when the system recommends to not switch from intravenous to oral antibiotics. This finding aligns with the large degree of cautious antibiotic prescribing reported by participants. Future research should carefully consider behavioural and social influences when developing technology for uncertain clinical decision making. More broadly, our results indicate that AI explanations are not frequently used and are unlikely to affect decision making outside of edge cases or conflicting scenarios. Regarding the translation of AI technologies into health care, we highlight that usability and clinical evidence are essential. As such, this AI CDSS has potential to support individualised prescribing decisions on switching from intravenous to oral antimicrobials, but further development of the user interface and prospective validation are required.

(MIMIC)-IV[19,20] (comprising 10 362 unique patient stays) and Imperial College Healthcare NHS Trust (ICHT) data (comprising 5610 unique patient stays), and had mean areas under the receiver operating characteristic curve of 0·80 (SD 0·01) and 0·79 (0·01), respectively, for the two datasets. Increasing the appropriate use of oral therapy and minimising intravenous administration is essential for preventing health-care-associated infections and promoting antimicrobial stewardship.[21–23] Oral antibiotics are often as effective as those administered intravenously, with added benefits including fewer side-effects, decreased nursing workload, lower costs, reduced climate impact, and improved patient comfort.[24–27] However, early intravenous-to-oral switching often remains underutilised owing to a lack of awareness and empowerment and a poor understanding of when individual patients are suitable for switching.[28–33]

We aimed to evaluate this AI-driven CDSS technology using a randomised, clinical vignette, web app-based study, alongside semistructured interviews and questionnaires, to understand the effect of the AI system and its explanations on decisions and participants' perspectives, and how such technology could fit into a clinical pathway.

Methods

Study design and participants

In this randomised, multimethod study, health-care professionals from across the UK who were regularly involved in antibiotic prescribing were recruited between April 23, 2024 and Aug 16, 2024, through personal networks and the general email list of the British Infection Association. Participants received an information sheet on the study before written informed consent was obtained. This study was approved by the Research Governance and Integrity Team at Imperial College London (Imperial College Research Ethics Committee reference 6880677).

The study was split into three phases: a pre-intervention semistructured interview, a clinical vignette experiment, and questionnaires. All phases of the study were completed in one 60–90-min online meeting. The clinical vignette experiment and questionnaires were conducted through a web app created using Django, HTML, and JavaScript. Before these two phases, clinicians' medical specialty, grade (level of seniority), age, sex, and familiarity with AI (collectively referred to as demographic information) were collected through the app. For statistical analysis, all demographic information was converted into categorical variables. Age was grouped by decade; sex (reported as free text by participants) into male and female; medical specialty into infectious diseases, microbiology, pharmacist, and other; and grade into consultant and training grade doctor. Familiarity with AI was assessed on a 5-point Likert scale. Any missing values were imputed with the median.

Semistructured interviews

The objective of the semistructured interviews was to understand clinicians' current experience of and role in antimicrobial prescribing, as well as their perception of AI. A topic guide (appendix p 5) containing example questions was developed to ensure all relevant points were covered during the discussions. Interviews were recorded and transcribed. To convert the raw qualitative data into quantitative data, we used a content analysis coding process.[34] A codebook was deductively developed on the basis of the topic guide to facilitate the coding process, with inductive processes to adapt and adjust codes as needed during analysis. Research over the past 3 years has highlighted the ability of large language models to assist with qualitative research.[35–37] To this extent, we used GPT-4o (released May 13, 2024) to fill in the code book for each transcript. Coding each transcript involved assigning a binary value (1 or 0) to each code depending on whether the participant mentioned that particular aspect during the interview or not. To ensure that the large language model coding reflected the researchers' perspectives, prompt engineering was initially used to improve the Cohen's κ agreement score between the large language model output and a manually coded transcript. Through this process, Cohen's κ was increased from 0·38 (fair agreement) to 0·72, which is considered substantial agreement.[38] The final prompt (appendix p 6) involves few-shot in-context learning[39] as well as providing the large language model with the codebook. In addition to prompt engineering, each comma-separated value output produced by the large language model was cross-checked by a member of the research team (WJB or TMR) against the original transcript to ensure consistency and prevent incorrect coding. Frequencies were calculated for analysis, and χ^2 (with Yates' correction) and Fisher's exact tests were used to assess differences in perspectives across participant demographics, with the Benjamini–Hochberg correction applied to control for multiple comparisons.

See Online for appendix

For the **codebook** see https://github.com/WilliamBolton/AI_IVOS_EVAL

Clinical vignette experiment

The objective of the clinical vignette experiment was to establish if and how an AI CDSS could influence decision making around the switch from intravenous to oral antibiotics. The experiment comprised 12 case vignettes, each involving a patient currently receiving intravenous antibiotics (range of treatment duration 30–127 h). Six were simulated cases generated to be representative of patients in ICHT, based on the empirical distributions observed in the dataset (hereafter denoted ICHT cases). The Imperial Clinical Analytics, Research and Evaluation (iCARE) environment was used to access de-identified, real-world clinical data for individuals with an inpatient encounter at ICHT. The design of these cases was informed by expert clinicians to focus on scenarios that they regularly encounter and to investigate how the AI tool influences decision making. The other six case vignettes were real patient cases randomly extracted from MIMIC-IV, a large, publicly available, de-identified, real-world clinical dataset (hereafter denoted MIMIC cases).[19,20] Details on the case vignettes are provided in table 1. The ground truth for

	Intervention		AI CDSS recommendation	Guideline CDSS recommendation	Ground truth	Data	Group
	Group A	Group B					
Patient 1	CDSS	SOC	Don't switch	Don't switch	Don't switch	ICHT	AI correct
Patient 2	SOC	CDSS	Switch	Switch	Switch	ICHT	AI correct
Patient 3	SOC	CDSS	Switch	Don't switch	Don't switch	ICHT	AI incorrect
Patient 4	CDSS	SOC	Don't switch	Switch	Switch	ICHT	AI incorrect
Patient 5	SOC	CDSS	Switch	Don't switch	..	ICHT	Uncertain
Patient 6	CDSS	SOC	Don't switch	Switch	..	ICHT	Uncertain
Patient 7	CDSS	SOC	Don't switch	Don't switch	Didn't switch	MIMIC-IV	AI correct
Patient 8	SOC	CDSS	Switch	Switch	Switched	MIMIC-IV	AI correct
Patient 9	SOC	CDSS	Potentially switch	Switch	Didn't switch	MIMIC-IV	Uncertain
Patient 10	CDSS	SOC	Don't switch	Switch	Switched	MIMIC-IV	AI incorrect
Patient 11	CDSS	SOC	Switch	Don't switch	Didn't switch	MIMIC-IV	AI incorrect
Patient 12	SOC	CDSS	Potentially switch	Don't switch	Switched	MIMIC-IV	Uncertain

For patient 5 and patient 6, a ground truth could not be defined. AI=artificial intelligence. CDSS=clinical decision support system. ICHT=Imperial College Healthcare NHS Trust. MIMIC=Medical Information Mart for Intensive Care. SOC=standard of care.

Table 1: **Details of the case vignettes**

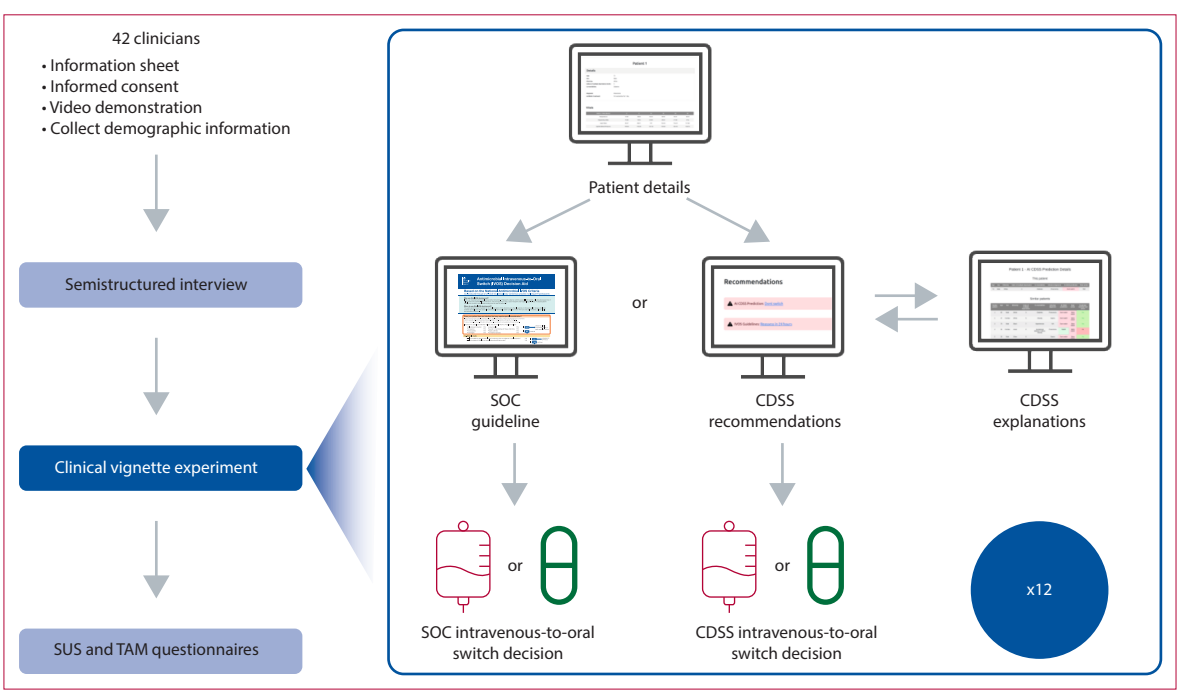

Figure 1: **Experimental protocol**

CDSS=clinical decision support system. SOC=standard of care. SUS=system usability scale. TAM=technology acceptance model.

MIMIC cases was defined as the real action that was taken; for simulated ICHT cases, it was decided by expert clinical collaborators. For each of the ICHT and MIMIC cases, we specified that two AI CDSS switch predictions would be correct relative to the ground truth, two would be incorrect, and two were uncertain. For each group consisting of two cases, the AI model recommended switching the antibiotic route in one case, whereas the recommendation for the other case was to not switch. All case vignettes were completed in the same order by all participants.

An illustration of the study design is shown in figure 1, with example screenshots of case vignettes from the web app shown in the appendix (p 4). For each case vignette, participants were asked to decide whether the patient in question was suitable for switching from intravenous to oral antibiotic treatment. The participants were given a binary switch or don't switch selection box as well as a free-text box to explain their reasoning or add any clarifications. Although this binary option does not encompass all of the possible permutations of clinical care seen for real patients, it enables us to focus this study on intravenous-to-oral antibiotic switching and disentangle the uncertainty in decision making on this subject, which is frequently seen. Participants were randomly assigned to one of two groups

(A or B) by the web app according to the order in which they began the experiment; the group to which a participant was assigned influenced what they viewed and had access to for a particular case vignette (table 1). Standard of care (SOC) information acted as our control. Here, participants could view the patient data (through a table and dynamic graph) and an embedded PDF of the UK Health Security Agency (UKHSA) antimicrobial intravenous-to-oral switch decision aid,[40] and they had a link to the associated UKHSA website.[41] Providing decision support acted as the intervention. The CDSS participants had access to all the information available in SOC, as well as the CDSS recommendations (from an AI model and guideline-based system using the UKHSA intravenous-to-oral switch criteria, appendix pp 2–3) and their explanations (such as similar patients, feature importance, and free-text insights). Explanations were provided on a separate web page through a clickable link to prevent information overload[42,43] and to ensure that we could reliably track when participants wanted to view AI explanations. For each group (A or B), each case vignette was assigned to either the SOC or CDSS intervention (table 1). The two groups were opposites of each other in terms of the intervention for a particular case vignette, with each group having six CDSS interventions (three for ICHT cases and three for MIMIC cases) and six SOC interventions (three for ICHT cases and three for MIMIC cases); this design ensured that the amount of data collected for each intervention across vignettes was equal and enabled estimation of between-intervention variability by controlling for the case. Data collected for each participant on each case vignette were intervention assigned, binary switch choice (switch or don't switch), free text explanation, mouse hover time on each element, URL clicks (eg, to see AI explanations), and time elapsed.

We used χ^2 (with Yates' correction) and Fisher's exact tests to test for significant differences in the switch choices between participants who were assigned the SOC versus the CDSS intervention for each case vignette. Subgroup analyses were conducted using these tests, with Benjamini–Hochberg correction applied to control the false discovery rate across demographic groups and questionnaire responses. Binary logistic regression, with the intervention as the independent variable and decision as the dependent variable, quantified the effect of the intervention on participants' switch choice for each case vignette. We used a generalised linear mixed-effects model with a binomial distribution and logit link function to examine the effect of the intervention across all clinicians' switching decisions. Random intercepts were incorporated for each case vignette and participant to account for between-patient variability and individual differences in decision making. In addition, a random slope for the intervention within clinicians was included to assess individual differences in response to the CDSS intervention.

Shannon entropy was calculated for each case vignette to assess the decision diversity across participants. Hover durations over various components within the experiment were aggregated, and regular expressions was used to extract particular terms from the free-text switch decision explanations, to act as a proxy for which aspects were having a greater influence on clinicians' decision making. Click-through rates for the AI and guideline CDSS explanation pages were summed, and Fisher's exact test was used to assess statistical differences on a per-patient level. Time spent on each case vignette was combined both across all participants and according to the intervention to which participants were assigned, and the mean and SD were calculated. For both the explainability click-through rate and duration, we used a paired t test (after meeting parametric requirements) to identify any differences in means. All statistical analyses used an α of 0·05. Overall, this study design and analysis plan enabled us to understand the effects of the AI system and its explanations and to test nuances such as whether health-care professionals can detect and ignore when the AI recommendations are incorrect.

Questionnaires

The final part of the study involved participants completing the system usability scale (SUS)[44] and technology acceptance model (TAM)[45] questionnaires. The purpose of the SUS survey is to assess participants' satisfaction with the application and interface, whereas the more nuanced questions in the TAM survey were designed to understand clinicians' acceptance of AI technology and opinions of the AI CDSS. The questions in both surveys used a 5-point Likert scale.[46] The SUS survey was analysed to return a SUS score; 68 is average and any score above this is considered good.[44] The TAM survey contained 28 questions and was developed on the basis of the TAM 2 methodology.[45,47,48] Some questions were assigned to categories common to the TAM framework: 12 for perceived usefulness, two for perceived ease of use (only two were included here owing to the SUS survey assessing the same concept), two for self-efficacy, two for subjective norm, and three for behaviour change; others were independent questions designed to assess a particular aspect from TAM or to evaluate a particular perspective or the influence of AI for antimicrobial decision making. At the end of the surveys, a free-text box was used to gather any final perspectives that the participants wanted to share. The internal consistency of TAM categories was measured with Cronbach's α.[49] ANOVA and Kruskal–Wallis tests, with multiple comparisons correction, were used to assess statistical differences in SUS scores and TAM responses across demographic groups. The choice of test depended on whether the data met parametric assumptions of homogeneity of variance (assessed using Levene's test) and normality (assessed using Shapiro–Wilk's test).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between April 23, 2024, and Aug 16, 2024, 59 clinicians were directly contacted or responded to outreach emails, with 42 clinicians from 23 hospitals in the UK participating in the study (table 2). The median age was 39 years (IQR 37–47), 19 (45%) participants were female and 23 (55%) were male, and 26 (62%) were consultants and 16 (38%) were training-grade doctors. 14 (33%) participants specialised in infectious diseases and 12 (29%) were microbiologists; the remaining participants were pharmacists (n=6 [14%]), or had no specialty or were from another area in medicine (hereafter termed other; n=10 [24%]). 17 (41%) participants were slightly familiar with AI and three (7%) were very familiar.

The results from coding the transcripts of the semi-structured interviews are presented in the appendix (p 10). 27 (64%) of 42 participants directly treated patients with antibiotics, 24 (57%) gave prescribing advice to other clinicians, and 14 (33%) were involved in shaping treatment or stewardship guidelines. There was disagreement between respondents on how individualised their current antimicrobial prescribing is, with one respondent noting, "We take patient-specific factors into account, but a lot of our choice is driven by what's in the guidelines." Behavioural and social aspects, including cautiousness (n=17 [40%]), resistance to change (n=10 [24%]), personal biases (n=8 [19%]), and the perception of risk towards long-term resistance (n=10 [24%]), were reported as influencing prescribing decisions. Participants reported using a range of technologies to support prescribing, including digital guidelines[50] (n=18 [43%]). However, 13 (31%) of 42 participants noted they had not used any technology to support clinical decision making. Most participants (n=26 [62%]) were optimistic about the use of AI, with clinical evidence (n=29 [69%]) on the amount of antimicrobial use (n=16 [38%]) and a range of patient outcomes, such as length of hospital stay (n=10 [24%]), being the biggest driver behind trust. Beyond optimising prescribing (n=19 [45%]), many participants noted the potential time (n=14 [33%]) and economic (n=8 [19%]) benefits of AI tools. Of the 42 participants, 14 (33%) and 15 (36%) mentioned that some form of explainability or understanding of the input features, respectively, would be necessary. Trust is also built over time through experience with a system (n=4 [10%]),[51] with participants noting that a shadow mode (n=6 [14%]), in which the tool is first run in the background without influencing decisions, and training (n=5 [12%]) could be useful. Ease of use was commonly cited as a factor influencing whether such an AI system is used (n=22 [52%]) and several barriers to adoption were highlighted, including challenges with changing behaviour (n=21 [50%]), hospitals' technological infrastructure not being ready for AI (n=19 [45%]), and the uncertain nature of antimicrobial prescribing decisions (n=14 [33%]).

For all but one case vignette, no significant differences were observed in participants' switch choices based on the intervention they were assigned (table 3). Three vignettes (patients 7, 10, and 11) resulted in almost equal disagreement between clinicians, with a Shannon entropy of 1·00 (a higher value indicates greater diversity in decisions). The mean time to complete a case vignette was 148·5 s (SD 26·1); no significant difference in completion times was observed between the SOC (143·6 s [SD 34·8]) and CDSS (153·5 s [26·4]) interventions (p=0·31). Hover duration (appendix p 12) was greatest for the table containing clinical data, followed by the graph that displayed the same information, with eight (19%) participants directly adjusting the variables displayed in the graph or the axis ranges. 23 (55%) of 42 participants viewed the AI CDSS explanations at any time, but across all case vignettes for which decision support was available, explanations were accessed 9% of the time. Similarly, 12 (29%) of 42 participants viewed an explanation of the guideline CDSS, but across all examples, the link was clicked only 5% of the time. AI CDSS explanations were viewed significantly more than guideline CDSS explanations (p=0·0012). The percentage of participants who requested explanations for each case vignette, grouped by decision support, is shown in figure 2A. No differences were observed across demographics. The mean time looking at the AI and guideline CDSS explanations was 59·8 s (SD 66·1) and 64·5 s (49·4), respectively, which accounted for, on average, 24·3% and 25·5% of the time

	Number of participants (N=42)
Grade	
Consultant	26 (62%)
Training-grade doctor	16 (38%)
Medical specialty	
Infectious diseases	14 (33%)
Microbiology	12 (29%)
Pharmacist	6 (14%)
Other	10 (24%)
Sex	
Male	23 (55%)
Female	19 (45%)
Age	
20–29 years	4 (105%)
30–39 years	17 (40%)
40–49 years	12 (29%)
50–59 years	6 (14%)
60–69 years	2 (5%)
Prefer not to say	1 (2%)
Median (IQR), years	39 (37–47)
Familiarity with AI	
Not familiar	9 (21%)
Slightly familiar	17 (40%)
Moderately familiar	9 (21%)
Very familiar	3 (7%)
No response	4 (10%)

Data are n (%) unless otherwise stated. AI=artificial intelligence.

Table 2: **Demographic characteristics of participants**

	Ground truth	Count				χ^2 (p value)	Fisher's odds ratio (p value)
		CDSS switch	CDSS don't switch	SOC switch	SOC don't switch		
Patient 1	Don't switch	0	21	0	21	0·00 (1·00)	NA
Patient 2	Switch	21	0	19	2	0·52 (0·47)	0·0 (0·49)
Patient 3	Don't switch	3	18	6	15	0·57 (0·45)	2·40 (0·49)
Patient 4	Switch	19	2	21	0	0·52 (0·47)	∞ (0·49)
Patient 5	··	21	0	19	2	0·52 (0·47)	0·0 (0·49)
Patient 6	··	17	4	21	0	2·49 (0·12)	∞ (0·11)
Patient 7	Didn't switch	6	15	16	5	7·73 (0·0054)	8·0 (0·0048)
Patient 8	Switched	13	8	12	9	0·0 (1·00)	0·82 (1·00)
Patient 9	Didn't switch	19	2	18	3	0·0 (1·00)	0·63 (1·00)
Patient 10	Switched	8	13	14	7	2·39 (0·12)	3·25 (0·12)
Patient 11	Didn't switch	11	10	11	10	0·00 (1·00)	1·00 (1·00)
Patient 12	Switched	6	15	9	12	0·41 (0·52)	1·88 (0·52)

CDSS=clinical decision support system. SOC=standard of care.

Table 3: **Switch decisions by intervention across case vignettes**

spent on the case. Five (12%) of 42 participants visited the UKHSA website at any time; however, across all case vignettes, the website was accessed only 1% of the time.

For one case vignette (patient 7, figure 2B) a significant difference was observed in the switch choice of participants according to whether they were assigned the CDSS or the SOC intervention (χ^2 [df=1, n=42] 7·73, p=0·0054; Fisher's exact test odds ratio [OR] 8·0 [95% CI 2·01–31·80]; p=0·0048). The majority of participants assigned to SOC chose to switch (16 [76%] of 21), whereas most of those assigned to the CDSS intervention chose to not switch (15 [71%] of 21). The recommendation of both the AI and guideline CDSS in this scenario was don't switch, which aligned with the ground truth action from the MIMIC-IV dataset. In this case, the binary logistic regression odds of switching from intravenous to oral antibiotics were significantly lower if the participant was assigned the CDSS intervention than if they were assigned SOC (OR 0·13, 95% CI 0·03–0·50; p=0·0031). Across all case vignettes and 504 observations, the generalised linear mixed-effects model showed similar results. The CDSS intervention was significantly associated with a decrease in the odds of switching compared with SOC (OR 0·60, 95% CI 0·42–0·86; p=0·0053). The model intercept was not significant (p=0·28) and patient-level variability was high (σ^2 4·75 [SD 2·40]), indicating heterogeneity in baseline switching behaviour across patients, which is expected given that different case vignettes had different ground truths (table 1). Clinician-level intercept variability was also present (σ^2 0·97 [0·99]), whereas the random slope variance for the intervention within clinicians was negligible (σ^2<0·01 [0·08]). Shifts in behaviour were also observed in some other vignettes, such as patient 6 and patient 10 (figure 2B); however, these differences were not significant. No significant differences were observed in subgroup analyses.

A score of 72·32 out of 100 (SD 8·79), which is considered good, was obtained for the SUS questionnaire (appendix p 13), with no significant differences in score observed across demographics. For the TAM questionnaire, Cronbach's α was 0·84 (good) for perceived usefulness, 0·52 (poor) for perceived ease of use, 0·51 (poor) for self-efficacy, 0·07 (unacceptable) for subjective norm, and 0·40 (unacceptable) for behaviour change.[49] Given the unacceptable scores for subjective norm and behaviour change, these questions were evaluated separately. All questions were scored on a scale from 1 to 5. The mean scores were 3·6 out of 5 (SD 0·31) for perceived usefulness, 3·8 out of 5 (0·20) for perceived ease of use, and 4·1 out of 5 (0·05) for self-efficacy (higher is better). The distribution of responses to some key questions in the TAM survey is shown in figure 3, and the complete set of results is in the appendix (p 14). Most (n=26 [62%]) of the 42 participants reported that they are comfortable with AI in health care, with their subjective norm (for either patients or people important to them) or perception of prestige having little effect on their use of AI technology. Although responses were variable, 21 (50%) of the 42 participants disagreed that their decisions were changed by the AI CDSS. However, their use of the AI CDSS was voluntary (n=42 [100%]), they were reassured when their clinical decision aligned with the AI CDSS recommendation (27 [77%] of 35), and they investigated further when their decision did not align (26 [74%] of 35). 27 (64%) of 42 participants agreed that the AI CDSS could be seamlessly integrated into their current clinical workflow, but there was disagreement on whether or not health-care institutions have the infrastructure to support such systems, which probably reflects the diverse range of hospitals represented in this study. Finally, 33 (79%) of 42 participants said they would intend to use the AI system if they had access to it, which was supported by a descriptive norm question in which 35 (83%) participants noted that some or most of their colleagues would use it.

Discussion

The objective of this study was to discern clinicians' perceptions of an AI CDSS focused on antibiotic prescribing and to understand if and how it can influence decision making regarding an intravenous-to-oral switch. For most

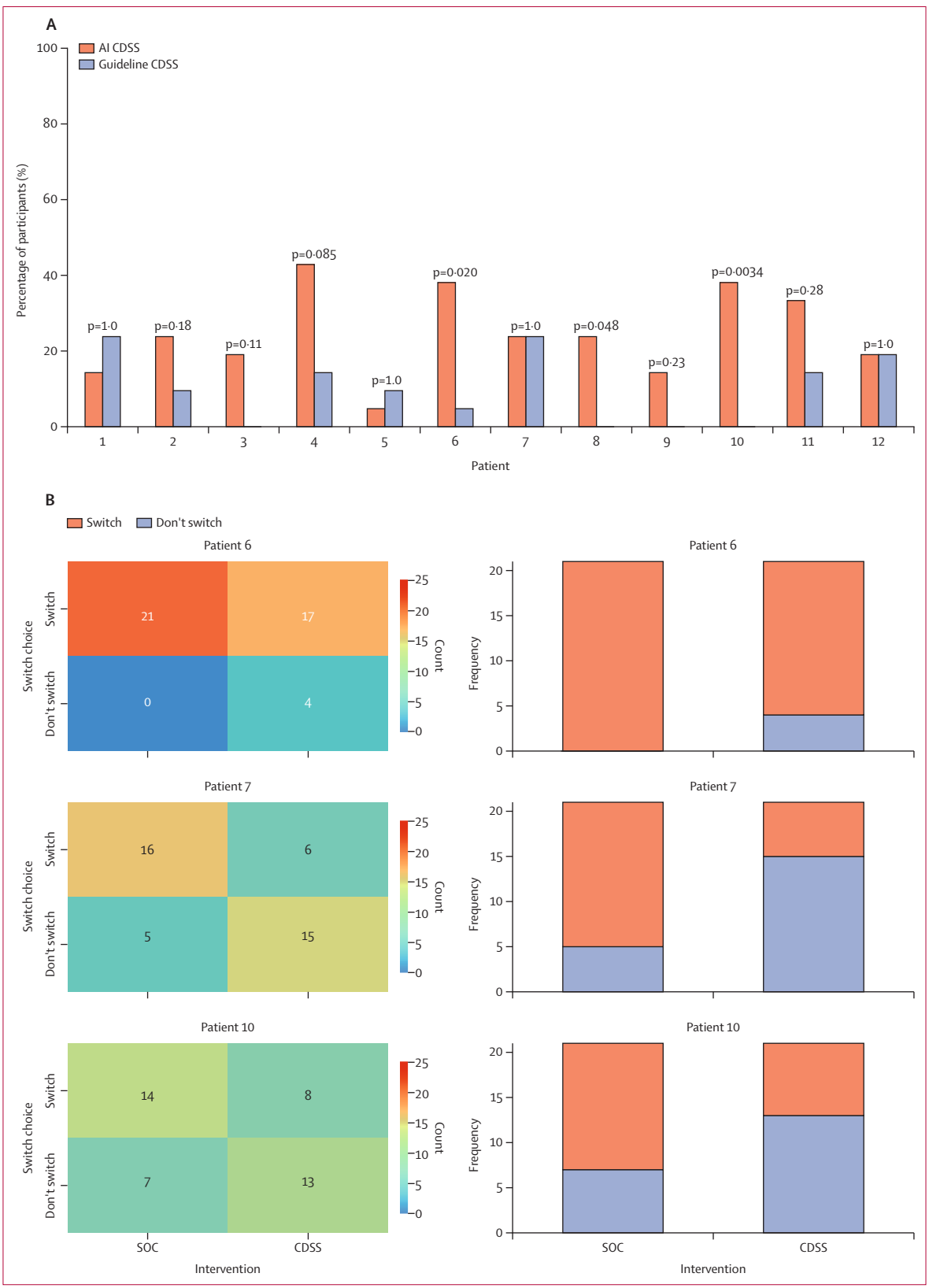

***Figure 2*: Utilisation of decision support explanations and effect of CDSS versus SOC intervention on intravenous-to-oral antibiotic switch decisions**
(A) Proportion of participants who clicked through for further explanation of the AI and guideline CDSS pages for each case vignette. p values shown are from Fisher's exact tests comparing click-through rates between the AI and guideline explanation pages on a per-patient level. (B) Participants' switch decisions for patients 6, 7, and 10 by intervention. Differences shown in (B) reflect the effect of decision support and not just explanations. AI=artificial intelligence. CDSS=clinical decision support system. SOC=standard of care.

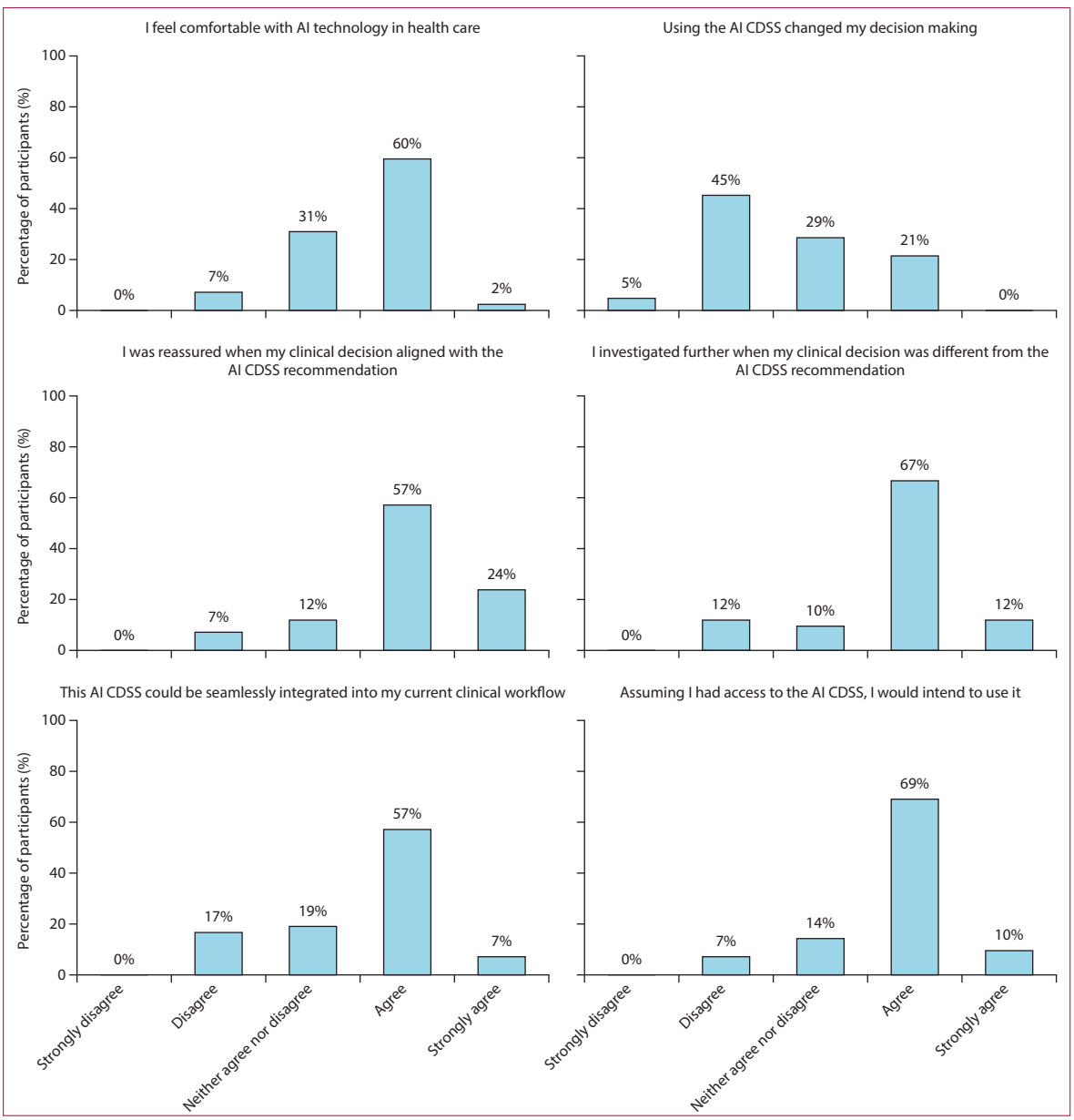

Figure 3: **Response distribution for key questions in the TAM questionnaire**
AI=artificial intelligence. CDSS=clinical decision support system. TAM=technology acceptance model.

case vignettes, no differences in decisions were observed between participants who had access to the CDSS intervention versus SOC control. This result is to be expected for patients 1 and 2, for whom the switch decision was designed to be obvious and the AI CDSS recommendation correct. For those cases in which the AI prediction was incorrect (patients 3, 4, 10, and 11), our findings highlight that clinicians could often correctly identify this and ignore the AI support, which contrasts with evidence from previous studies.[52–54] When shifts in decision making were observed, the AI CDSS recommended that switching would not be appropriate, and the change in decisions by participants who were assigned the CDSS intervention, relative to those who were assigned SOC, reflected this. This observation implies that clinicians are persuaded to take the option that is perceived as safer and not switch when the AI system supports this, regardless of their initial inclination, reflecting the culture of cautiousness (40%) and hesitancy to change (24%) reported by participants in interviews. The finding that AI explanations were used the most (by 43% of participants; figure 2B) for patient 4 also supports this argument: in this case vignette, the AI recommendation was don't switch, which was incorrect relative to the ground truth of switch. Almost all participants

correctly ignored this advice and continued with their decision to switch. However, almost half of the clinicians examined the AI explanation for this patient, compared with only 19% for patient 3, for whom the situation was reversed. This finding suggests that, even though switching was the appropriate decision in this case, the AI CDSS recommendation to not switch had a greater effect on participants' actions than did a recommendation to switch in the opposite scenario. This behaviour also aligns with a question from the TAM survey, in which participants reported that they investigated further when their clinical decision was different from the AI CDSS recommendation (74%). The results from the generalised linear mixed-effects model suggest that the effect of the CDSS was relatively consistent across participants, with the random slope variance for intervention within participants near to zero. Half of the participants reported that the AI CDSS did not change their decision making, but no statistical differences were observed in switch choices based on this factor, implying that self-reporting might not be a valid metric in this setting.[55] Results from the SUS and TAM questionnaires were promising, indicating that the AI CDSS was useful, usable, and appropriate, and could be integrated into current workflows. Participants were reassured when the AI recommendation aligned with their clinical decision, consistent with previous research showing that expert clinicians are more confident when given AI advice.[56] Overall, our results indicate that the AI CDSS was positively received, with decision patterns largely unchanged, and any influence directing clinicians towards not switching from intravenous to oral antibiotics.

The added value of including or excluding explanations for AI CDSSs has been previously studied, with numerous publications noting that explainable AI often has little effect.[51,54,55,57] In this research, we aimed to understand the extent to which AI explanations were used when available. Around half of the participants (55%) viewed the AI explanations at least once during the study. However, across all cases for which decision support was available, explanations were used only 9% of the time, with around a quarter of the duration of a case vignette being dedicated to looking at the AI explanation. Even when considering participants who were assigned to the CDSS intervention and chose to align with its recommendation for patient 7 (a vignette for which a significant difference in decision making was observed between participants assigned to CDSS and SOC), only 13% viewed the AI explanation, implying that it was not the persuasive nature of the explanation that changed their decision making, but rather the presence of the AI recommendation itself. No correlation was observed between participants' familiarity with AI and whether they were comfortable with AI in health care, which also indicates that understanding the technical details of an AI system is not a key driver behind adoption.[58] On the other hand, 33% of participants mentioned explainable AI in their interview (appendix p 10), with one participant noting: "If you don't explain it to somebody on the front line, they're going to get frustrated, and all they'll do is just start ignoring the system." However, time is precious when making clinical decisions, and the same participant noted "I don't have time to look up loads of other stuff", and that free-text explanations should just be incorporated directly into the patient's clinical notes. This lack of time to seek additional information is supported by the low proportion of participants who interacted with the dynamic graph displaying clinical data (19%) and the fact that, across all case vignettes, the UKHSA website was accessed only 1% of the time. Understanding an AI model is essential when developing CDSSs for health care to ensure models align with human intuition. At the point of use, however, our findings agree with previous studies indicating that explanations probably have little effect on the influence of an AI system in health care.[51,54,55,57] Evidence from this study suggests that this lack of effect of explanations might be driven by time constraints and increased cognitive load. As such, future work should look into how to effectively train users to interpret explanations while also emphasising simple methods, such as displaying similar patient cases and using large language models to summarise information into text. By aligning explanations with clinicians' preferences and time pressures, and by assessing the use of these explanations, we can help to ensure that AI CDSSs are used appropriately.

Bacterial AMR is a growing global threat. Estimated to cause around 1·14 million deaths annually,[59] AMR also has considerable financial implications, including a cost of more than £180 million each year to the UK's National Health Service.[60] To mitigate the increase in AMR, we must optimise antimicrobial use and extend the therapeutic life of existing drugs, which is known as antimicrobial stewardship. Prescribing practices and stewardship activities exhibit substantial global variability.[61] Despite considerable research efforts,[62,63] these decisions are inherently complex,[29] and what is best for an individual patient is often unknown, with one participant in our study saying: "I see very little evidence for an awful lot of what we do." We observed high clinician variability in intravenous-to-oral switch decisions regardless of the intervention to which the participant was assigned. Variability was particularly high for the MIMIC cases, which, excluding patient 9, all had a Shannon entropy of at least 0·94 (appendix p 11). This variability is probably due to the reduced clinical information available to participants in these cases; however, such situations are common in clinical practice,[10] highlighting a key barrier when applying AI to infectious disease. If there is insufficient scientific understanding or information to support optimal decision making, then heterogeneity in practice can prevent AI models trained on such data from converging effectively. AI outputs can then come into contradiction with a clinician's mental model of what is considered optimal care, leading to a breakdown in confidence and such systems being rejected by the user. One participant stated: "I think the challenge will come if it's suggesting you do something very opposed to your normal

practice", with the broader interview results showing that the key bottlenecks in translating such AI systems to the clinic are the uncertainty of antimicrobial prescribing decisions (33%) and the challenges involved in getting clinicians to change their existing prescribing behaviour (50%). Indeed, defining a ground truth for some of the case vignettes in this study was not possible because of this uncertainty, highlighting that obtaining high-quality labels in such scenarios is non-trivial. The ability of infectious diseases to affect others and the development of AMR also raises challenges, including balancing the needs of an individual with those of wider society[11] (24%) and updating the system temporally and geographically to match changes in pathogens and their resistance patterns. How such updates can be addressed while still passing regulatory requirements is an unresolved question. Ultimately, even if uncertainty exists, tools that assist with antimicrobial prescribing and are supported by evidence can still provide a benefit to clinicians and patients and help with tackling AMR. However, infectious diseases pose unique hurdles for AI CDSSs, with one participant stating: "The layers of complexity on an individual basis is why microbiology as a service exists."

Behavioural, social, and cultural factors have an important role in antimicrobial prescribing.[13,64–66] Participants noted that prescribers are often biased towards particular antimicrobial agents on the basis of their lived experience (19%), which can lead to cautious prescribing (40%) that unnecessarily extends the duration of intravenous antibiotics. The influence of these factors varies by specialty and seniority, with training-grade doctors being heavily influenced by their more senior peers (24%). Participants mentioned that the frequent rotation of junior doctors and the presence of temporary staff mean that stewardship education must be continuously restarted. Another factor that influences adherence to stewardship practices is resources, including the presence of pharmacists and microbiologists on hospital wards. When present, such experts can provide regular advice and steer clinicians towards optimal prescribing. Unfortunately, in resource-constrained environments, such experts are frequently absent, leading to the potential for an AI system to augment this support through readily available technology, including mobile phones.[67] However, the use of new systems in this way could potentially be overwhelming from a user perspective when health-care systems are stretched. To ensure adoption, data showing the safety and benefit of an AI CDSS is necessary (69%), with clinicians wanting some understanding of how the system was devised and the information it uses (36%). Regarding outcome measures, patient outcomes were considered most important,[14] but participants noted that showing a benefit in, for example, mortality or treatment failure rate would be difficult for a prescribing and stewardship-focused CDSS. Instead, demonstrating non-inferiority in terms of patient outcomes and a benefit elsewhere, such as reduced antimicrobial consumption (38%) and resistance, or cost savings (19%) through clinician time (33%) or bed days saved (24%), was seen as an appropriate approach. We found little evidence that other stakeholders (eg, colleagues and patients) have a large influence on a clinician's perception of an AI system (appendix p 10), which is consistent with some previous research.[68] However, it is likely that, on a wider scale, an organisation's ethos, support, training (12%), and social norm feedback would be essential for driving behaviour change.[8,69–71] Indeed, the use of such systems is supported by updates in 2023 and 2024 to the UK's regulations regarding software as a medical device, which ensure that when and how the systems are used are clearly defined.[72,73] Finally, as with any software, maintenance and contingency plans for when things go wrong are important. Taken together, setting up the appropriate support around an AI CDSS and gathering robust data is important to drive a behaviour change in traditionally heterogeneous and conservative prescribing practices.

Limitations exist for this study. First, although a diverse population of clinicians participated, our recruitment primarily relied on individuals responding to our emails. This strategy inherently limited the size of our study, meaning that small-to-medium effects could not be detected, and the participant population was biased towards those who are interested in the research and AI technology. However, interest in the topic did not necessarily mean that the participants were familiar with AI. Second, the clinical vignette experiment in this research was conducted through a web app, which, despite using simulated or real patient cases, is inherently different from a clinical setting. Concordant with this is the fact that, despite each case containing a large amount of infection-relevant data (as noted by participants), not all the information often present in real clinical practice could be included owing to dataset limitations. Accordingly, several participants noted that they would request additional information to confirm their decisions. Finally, coding the interview transcripts could have oversimplified the data, and participants' internal decision-making processes were not collected in the clinical vignette experiment. As such, when analysing the results, some level of inference must be applied, which is possible by triangulating findings across the mixed methods used in this study. However, future work could use more formal approaches, such as so-called thinking-aloud methods[74] or eye tracking.[75]

This research highlights key findings that can be broadly applied to AI CDSSs in health care. In particular, these systems can often have unexpected effects on decision making; AI explanations are likely to be underused in clinical practice; and non-interoperable, heterogeneous infrastructure is a barrier to implementation. To drive clinical adoption of AI CDSSs, developers and designers should look to maximise trust, which is most often built slowly over time through robust clinical evidence, experience, network effects, and by following frameworks such as FUTURE-AI.[76,77] To be accepted, tools must also be easy to use, and the co-design of systems and training has been

proposed to support appropriate user experiences and use.[78,79] Understanding what behaviour change is needed and how the tool fits into clinical workflows is essential to ensure impact and appropriate patient and clinician engagement. Ideally, decision support systems are integrated into electronic health records such that minimal input or behaviour change is required from end users; however, such integration can often be complex and infeasible, leading to separate apps with simple and attractive user interfaces being an alternative option. On the basis of this research, it seems there are two use cases in which this CDSS focused on intravenous-to-oral antimicrobial switching could provide the most utility. First, the tool can empower clinicians who are not experts in antimicrobial stewardship, particularly in settings in which staff have little expertise in infection, with one participant saying: "If junior doctors were able to be more independent prescribers, as a result of this tool, then that would be a good thing." Second, the tool can serve as a risk-stratification system for pharmacists and microbiologists, helping them to focus their time on treating patients with the most complex conditions. Results show that participants had a positive perception of this system; however, while the system influenced decision making regarding the intravenous-to-oral switch to some extent, it did not lead to an increase in switching to oral antimicrobials. Further work is therefore needed to understand how AI CDSSs interact with the existing behavioural and social influences on antimicrobial prescribing to effectively impact early intravenous-to-oral switching and fully realise the benefits of oral antimicrobials.

This AI CDSS has the potential to support optimal antimicrobial prescribing and drive the cultural and behavioural change needed at a system level to tackle AMR. Results show that the AI CDSS was well received by participants and can affect antimicrobial prescribing decisions in some situations, in particular when it recommends not switching from intravenous to oral antibiotics, which aligns with the cautious behaviour reported by participants. However, prospective clinical trials are required to gather the necessary evidence on safety and benefits. AI systems hold promise in supporting uncertain clinical decisions. However, the great complexity of medicine necessitates research such as this to understand the effects, perception, and interaction with existing decision-making processes and social influences, to ultimately bridge the translation gap and benefit patients.

Contributors

WJB and TMR contributed to the concept and design of the study. WJB contributed to data acquisition and analysis. WJB and TMR contributed to the manuscript drafting, discussion of the results, and accessed, verified, and reviewed the raw data. All authors contributed to data interpretation, as well as final revisions of the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

TMR received honoraria from Sandoz (2020), Roche Diagnostics (2021), Pfizer (2022 and 2024), and bioMérieux (2020–25). MG has advised Pfizer, Menarini, Napp, and GSK. PG is a co-founder and director of ProtonDx. AH is on the Scientific Advisory Group for Emergencies Coronavirus Response working group on nosocomial transmission and the WHO Health Emergencies Program Ad-Hoc Advisory Panel of Infection Prevention and Control Experts for Preparedness, Readiness and Response to COVID-19; leads and received funding for the Centres for Antimicrobial Optimisation Network via the University of Liverpool and Imperial College London; is the Director of The Fleming Initiative, a partnership between Imperial College London, Imperial College Healthcare NHS Trust, and commercial and philanthropic partners including GSK, Cepheid, Optum, and LifeArc; was previously an executive committee member and past President of the International Society for Infectious Diseases, a board member for the Wellcome Surveillance and Epidemiology of Drug-resistant Infections Consortium, and chair of the Fleming Fund Technical Advisory Group (operated by the UK Department of Health and Social Care); and has patents in development (three related to diagnostic innovations related to point-of-care testing for therapeutic drug monitoring, two for point-of-care perinatal host factor response monitoring, and a non-invasive innovation for the application of photoplethysmography in the continuous monitoring of circulating volume in hospital settings). All other authors declare no competing interests.

Data sharing

The MIMIC-IV dataset is publicly available.[19] The dataset is accessible to those who are credentialed users on PhysioNet, have completed the required training, and signed the appropriate data use agreement. Data from ICHT can be accessed by contacting the iCARE team and completing the steps described at https://www.imperial.ac.uk/medicine/research-and-impact/groups/icare/icare-facility/information-for-researchers/. Data generated and analysed and the code used during this study are available at https://github.com/WilliamBolton/AI_IVOS_EVAL.

Acknowledgments

WJB was supported by the UK Research and Innovation Centre for Doctoral Training in AI for Healthcare (https://ai4health.io; grant number P/S023283/1) and the National Institute for Health and Care Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at Imperial College London, in partnership with the UK Health Security Agency (previously Public Health England) in collaboration with Imperial Healthcare Partners, the University of Cambridge, and the University of Warwick. AH is a National Institute for Health and Care Research Senior Investigator. We thank the Centre for Antimicrobial Optimisation at Imperial College London, funded by the UK Department for Health and Social Care, and the iCARE team at ICHT. This research was enabled by the iCARE environment and used the iCARE team and data resources. These entities were not involved in the study design; the collection, analysis, or interpretation of the data; the writing of this Article, or the decision to submit for publication. The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the National Institute for Health and Care Research, the UK Department of Health and Social Care, or the UK Health Security Agency.

References

1 Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ* 2023; **23:** 689.

2 Higgins D, Madai VI. From bit to bedside: a practical framework for artificial intelligence product development in healthcare. *Adv Intell Syst* 2020; **2:** 2000052.

3 McCradden MD, Stephenson EA, Anderson JA. Clinical research underlies ethical integration of healthcare artificial intelligence. *Nat Med* 2020; **26:** 1325–26.

4 Olaye IM, Seixas AA. The gap between AI and bedside: participatory workshop on the barriers to the integration, translation, and adoption of digital health care and AI startup technology into clinical practice. *J Med Internet Res* 2023; **25:** e32962.

5 Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. *Cureus* 2023; **15:** e46454.

6 Lambert SI, Madi M, Sopka S, et al. An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. *npj Digit Med* 2023; **6:** 111.

7 Peiffer-Smadja N, Rawson TM, Ahmad R, et al. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clin Microbiol Infect* 2020; **26:** 584–95.

8 Tokgöz P, Krayter S, Hafner J, Dockweiler C. Decision support systems for antibiotic prescription in hospitals: a survey with hospital managers on factors for implementation. *BMC Med Inform Decis Mak* 2024; **24:** 96.

9 Mello-Thoms C, Mello CAB. Clinical applications of artificial intelligence in radiology. *Br J Radiol* 2023; **96:** 20221031.

10 Wojcik G, Ring N, McCulloch C, Willis DS, Williams B, Kydonaki K. Understanding the complexities of antibiotic prescribing behaviour in acute hospitals: a systematic review and meta-ethnography. *Arch Public Health* 2021; **79:** 134.

11 Bolton WJ, Badea C, Georgiou P, et al. Developing moral AI to support decision-making about antimicrobial use. *Nat Mach Intell* 2022; **4:** 912–15.

12 Tarrant C, Krockow EM. Antibiotic overuse: managing uncertainty and mitigating against overtreatment. *BMJ Qual Saf* 2022; **31:** 163–67.

13 Poluektova O, Robertson DA, Rafferty A, Cunney R, Lunn PD. A scoping review and behavioural analysis of factors underlying overuse of antimicrobials. *JAC Antimicrob Resist* 2023; **5:** dlad043.

14 Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; **17:** 195.

15 Sendak MP, D'Arcy J, Kashyap S. A path for translation of machine learning products into healthcare delivery. *EMJ Innov* 2020. https://doi.org/10.33590/emjinnov/19-00172.

16 Vasey B, Nagendran M, Campbell B. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med* 2022; **28:** 924–33.

17 White NM, Cater HE, Kularatna S, et al. Evaluating the costs and consequences of computerized clinical decision support systems in hospitals: a scoping review and recommendations for future practice. *J Am Med Inform Assoc* 2023; **30:** 1205–18.

18 Bolton WJ, Wilson R, Gilchrist M, Georgiou P, Holmes H, Rawson TM. Personalising intravenous to oral antibiotic switch decision making through fair interpretable machine learning. *Nat Commun* 2024; **15:** 506.

19 Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV. June 12, 2022. https://physionet.org/content/mimiciv/2.0/ (accessed Aug 30, 2024).

20 Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000; **101:** E215–20.

21 UK Health Security Agency. Antimicrobial stewardship: start smart then focus. March 25, 2015. https://www.gov.uk/government/publications/antimicrobial-stewardship-start-smart-then-focus (accessed Aug 30, 2024).

22 Davar K, Clark D, Centor RM, et al. Can the future of ID escape the inertial dogma of its past? The exemplars of shorter is better and oral is the new IV. *Open Forum Infect Dis* 2022; **10:** ofac706.

23 Sendi P, Nelson SB, Soriano A, Spellberg B. Early switch from intravenous to oral anti-microbial therapy in infectious diseases. *Clin Microbiol Infect* 2023; **29:** 1117–19.

24 McMeekin N, Geue C, Briggs A, et al. Cost-effectiveness of oral versus intravenous antibiotics (OVIVA) in patients with bone and joint infection: evidence from a non-inferiority trial. *Wellcome Open Res* 2019; **4:** 108.

25 Wald-Dickler N, Holtom PD, Phillips MC, et al. Oral is the new IV. Challenging decades of blood and bone infection dogma: a systematic review. *Am J Med* 2022; **135:** 369–79.e1.

26 Gilchrist M, Barr D, Drummond F, et al. Outpatient parenteral antimicrobial therapy (OPAT) in the UK: findings from the BSAC National Outcomes Registry (2015–19). *J Antimicrob Chemother* 2022; **77:** 1481–90.

27 Tan M, Powell N, Hearsey D. P35 impact of IV to oral antibiotic switch optimization on nurse time and plastic waste at the Royal Cornwall Hospital Trust. *JAC Antimicrob Resist* 2024; **6** (Suppl 2): dlae136.039.

28 Shrayteh ZM, Rahal MK, Malaeb DN. Practice of switch from intravenous to oral antibiotics. *SpringerPlus* 2014; **3:** 717.

29 Rawson TM, Charani E, Prockter Moore LS, et al. Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study. *BMC Med* 2016; **14:** 208.

30 Hospenthal DR, Waters CD, Beekmann SE, Polgreen PM. Practice patterns of infectious diseases physicians in transitioning from intravenous to oral therapy in patients with bacteremia. *Open Forum Infect Dis* 2019; **7:** ofz386.

31 Buis DTP, Prins JM, Betica-Radic L, et al. Current clinical practice in antibiotic treatment of *Staphylococcus aureus* bacteraemia: results from a survey in five European countries. *J Antimicrob Chemother* 2022; **77:** 2827–34.

32 Pandolfo AM, Horne R, Jani Y, et al. Understanding decisions about antibiotic prescribing in ICU: an application of the Necessity Concerns Framework. *BMJ Qual Saf* 2022; **31:** 199–210.

33 Hamilton RA, Williams N, Ashton C, et al. Nurses' attitudes, behaviours, and enablers of intravenous to oral switching (IVOS) of antibiotics: a mixed-methods survey of nursing staff in secondary care hospitals across the Midlands region of England. *J Hosp Infect* 2024; **150:** 9–16.

34 Erlingsson C, Brysiewicz P. A hands-on guide to doing content analysis. *Afr J Emerg Med* 2017; **7:** 93–99.

35 Xiao Z, Yuan X, Liao QV, Abdelghani R, Oudeyer P-Y. Supporting qualitative analysis with large language models: combining codebook with GPT-3 for deductive coding. 28th International Conference on Intelligent User Interfaces. March 27–31, 2023. https://doi.org/10.1145/3581754.3584136 (75–78).

36 Dunivin ZO. Scalable qualitative coding with LLMs: chain-of-thought reasoning matches human performance in some hermeneutic tasks. *arXiv* 2024; published online Feb 12. https://doi.org/10.48550/arXiv.2401.15170 (preprint).

37 Tai RH, Bentley LR, Monteith BG. An examination of the use of large language models to aid analysis of textual data. *Int J Qual Methods* 2024; **33**. https://doi.org/10.1177/16094069241231168.

38 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33:** 159–74.

39 Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst* 2020; **33:** 1877–901.

40 UK Health Security Agency. Antimicrobial intravenous-to-oral switch (IVOS) decision aid. January, 2023. https://assets.publishing.service.gov.uk/media/63da4c5de90e0773d99d59d0/Final_IV_to_Oral_Switch_Decision_Aid_based_on_NationalCriteria_UKHSA.pdf (accessed Aug 30, 2024).

41 UK Health Security Agency. National antimicrobial intravenous-to-oral switch (IVOS) criteria for early switch. January, 2023. https://www.gov.uk/government/publications/antimicrobial-intravenous-to-oral-switch-criteria-for-early-switch/national-antimicrobial-intravenous-to-oral-switch-ivos-criteria-for-early-switch (accessed Aug 30, 2024).

42 Laker LF, Froehle CM, Windeler JB, Lindsell CJ. Quality and efficiency of the clinical decision-making process: information overload and emphasis framing. *Prod Oper Manag* 2018; **27:** 2213–25.

43 Phillips-Wren G, Adya M. Decision making under stress: the role of information overload, time pressure, complexity, and uncertainty. *J Decis Syst* 2020; **29** (supp 1): 213–25.

44 Lewis JR. The system usability scale: past, present, and future. *Int J Hum Comput Interact* 2018; **34:** 577–90.

45 Holden RJ, Karsh B-T. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010; **43:** 159–72.

46 Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932; **22:** 55.

47 Ketikidis P, Dimitrovski T, Lazuras L, Bath PA. Acceptance of health information technology in health professionals: an application of the revised technology acceptance model. *Health Inform J* 2012; **18:** 124–34.

48 Buenestado D, Elorz J, Pérez-Yarza EG, et al. Evaluating acceptance and user experience of a guideline-based clinical decision support system execution platform. *J Med Syst* 2013; **37:** 9910.

49 Nunnally JC, Bernstein IH. Psychometric theory, 3rd edn. New York: McGraw–Hill, 1994.

50 Hand KS, Clancy B, Allen M, Mayes A, Patel Y, Latter SM. 'It makes life so much easier'—experiences of users of the MicroGuide™ smartphone app for improving antibiotic prescribing behaviour in UK hospitals: an interview study. *JAC Antimicrob Resist* 2021; **3:** dlab111.

51 Shafti A, Derks V, Kay H, Faisal AA. The response shift paradigm to quantify human trust in AI recommendations. *arXiv* 2022; published online Feb 16. https://doi.org/10.48550/arXiv.2202.08979 (preprint).

52 Gaube S, Suresh H, Raue M, et al. Do as AI say: susceptibility in deployment of clinical decision-aids. *npj Digit Med* 2021; **4:** 31.

53 Jacobs M, Pradier MF, McCoy TH Jr, Perlis RH, Doshi-Velez F, Gajos KZ. How machine-learning recommendations influence clinician treatment selections: the example of antidepressant selection. *Translat Psychiatry* 2021; **11:** 108.

54 Jabbour S, Fouhey D, Shepard S, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. *JAMA* 2023; **330:** 2275–84.

55 Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Quantifying the impact of AI recommendations with explanations on prescription decision making. *npj Digit Med* 2023; **6:** 206.

56 Gaube S, Suresh H, Raue M, et al. Non-task expert physicians benefit from correct explainable AI advice when reviewing X-rays. *Sci Rep* 2023; **13:** 1383.

57 Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 2021; **3:** e745–50.

58 Henry KE, Kornfield R, Sridharan A, et al. Human–machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *npj Digit Med* 2022; **5:** 97.

59 GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. English. *Lancet* 2024; **404:** 1199–226.

60 House of Commons. Health and Social Care Committee—oral evidence: antimicrobial resistance, HC 962. Sept 4, 2018. https://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/health-and-social-care-committee/antimicrobial-resistance/oral/88745.html (accessed Aug 30, 2024).

61 Pauwels I, Versporten A, Drapier N, Vlieghe E, Goossens H, the Global-PPS network. Hospital antibiotic prescribing patterns in adult patients according to the WHO Access, Watch and Reserve classification (AWaRe): results from a worldwide point prevalence survey in 69 countries. *J Antimicrob Chemother* 2011; **76:** 1614–24.

62 Charani E, McKee M, Ahmad R, et al. Optimising antimicrobial use in humans—review of current evidence and an interdisciplinary consensus on key priorities for research. *Lancet Reg Health Eur* 2021; **29:** 100161.

63 Rawson TM, Wilson RC, O'Hare D, et al. Optimizing antimicrobial use: challenges, advances and opportunities. *Nat Rev Microbiol* 2021; **19:** 747–58.

64 Lorencatto F, Charani E, Sevdalis N, Tarrant C, Davey P. Driving sustainable change in antimicrobial prescribing practice: how can social and behavioural sciences help? *J Antimicrob Chemother* 2018; **73:** 2613–24.

65 Donisi V, Sibani M, Carrara E, et al. Emotional, cognitive and social factors of antimicrobial prescribing: can antimicrobial stewardship intervention be effective without addressing psycho-social factors? *J Antimicrob Chemother* 2019; **74:** 2844–47.

66 Dionisio F, Baquero F, Fuertes M. Psychological and cultural factors influencing antibiotic prescription. *Trends Microbiol* 2023; **31:** 559–70.

67 Osei E, Mashamba-Thompson TP. Mobile health applications for disease screening and treatment support in low-and middle-income countries: a narrative review. *Heliyon* 2021; **7:** e06639.

68 Nadri H, Rahimi B, Lotfnezhad Afshar H, Samadbeik M, Garavand A. Factors affecting acceptance of hospital information systems based on extended technology acceptance model: a case study in three paraclinical departments. *Appl Clin Inform* 2018; **9:** 238–47.

69 Huang Z, George MM, Tan Y-R, et al. Are physicians ready for precision antibiotic prescribing? A qualitative analysis of the acceptance of artificial intelligence-enabled clinical decision support systems in India and Singapore. *J Glob Antimicrob Resist* 2023; **35:** 76–85.

70 Zeng Y, Shi L, Liu C, et al. Effects of social norm feedback on antibiotic prescribing and its characteristics in behaviour change techniques: a mixed-methods systematic review. *Lancet Infect Dis* 2023; **23:** e175–84.

71 Tokgöz P, Hafner J, Dockweiler C. Factors influencing the implementation of decision support systems for antibiotic prescription in hospitals: a systematic review. *BMC Med Inform Decis Mak* 2023; **23:** 27.

72 Medicines & Healthcare products Regulatory Agency. Crafting an intended purpose in the context of software as a medical device (SaMD). March 22, 2023. https://www.gov.uk/government/publications/crafting-an-intended-purpose-in-the-context-of-software-as-a-medical-device-samd/crafting-an-intended-purpose-in-the-context-of-software-as-a-medical-device-samd (accessed Aug 30, 2024).

73 Medicines & Healthcare products Regulatory Agency. Impact of AI on the regulation of medical products. April 30, 2024. https://www.gov.uk/government/publications/impact-of-ai-on-the-regulation-of-medical-products/impact-of-ai-on-the-regulation-of-medical-products (accessed Aug 30, 2024).

74 Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? *Qual Res Sport Exerc Health* 2017; **9:** 514–31.

75 Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Eye tracking insights into physician behaviour with safe and unsafe explainable AI recommendations. *npj Digit Med* 2024; **7:** 202.

76 Jones C, Thornton J, Wyatt JC. Enhancing trust in clinical decision support systems: a framework for developers. *BMJ Health Care Inform* 2021; **28:** e100247.

77 Lekadir K, Frangi AF, Porras AR, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ* 2025; **388:** e081554.

78 Rawson TM, Ming D, Gowers SA, et al. Public acceptability of computer-controlled antibiotic management: an exploration of automated dosing and opportunities for implementation. *J Infect* 2019; **78:** 75–86.

79 Wang L, Zhang Z, Wang D, et al. Human-centered design and evaluation of AI-empowered clinical decision support systems: a systematic review. *Front Comput Sci* 2023; **5**. https://doi.org/10.3389/fcomp.2023.1187299.