Implementation of artificial intelligence-based decision support systems for antibiotic prescribing in hospitals: a Delphi study

Pinar Tokgöz*, Joanna Albrecht and Christoph Dockweiler

Faculty of Arts and Humanities, Professorship of Digital Public Health, University of Siegen, Siegen, Germany

Introduction: Numerous initiatives against antimicrobial resistance have been initiated in recent years. Decision support systems (DSSs) based on artificial intelligence (AI) provide new opportunities for automating antibiotic therapy in hospitals. While AI-based DSSs may improve antimicrobial use and patient outcomes and reduce healthcare costs, the challenges associated with their implementation, optimization, and adoption cannot be ignored.

Methods: A Delphi study was conducted to investigate factors influencing the implementation of AI-based DSSs in the hospital setting.

Results: The study included 36 experts with perspectives on the hospital setting and DSS development. A consensus was reached on the importance of 34 factors and the ranking as well as assessment of current realization of implementation factors revealed important starting points for implementation strategies.

Discussion: The study results indicate that whilst there are multiple factors of importance in DSS implementation, some factors, as e.g., promoting application- and user-orientated development of DSSs, establishing user-friendly organizational structures, and fulfilling demands of trust, transparency, and responsibility through sensitization and education on organizational but also legal level should gain more attention. In addition, two factors did not reach a consensus in terms of importance, indicating that it may not be practical to consider all factors of importance when implementing AI-based DSSs in the hospital setting.

KEYWORDS

antibiotic prescribing, artificial intelligence, decision support, clinical practice, implementation

1 Introduction

Combatting antimicrobial resistance is widely regarded as a priority area in public health and several strategies haven been developed in response (1). One of the main causes of this problem is the over- or misuse of antibiotics (2). For this reason, antimicrobial stewardship programs (ASPs) have been established to optimize antimicrobial usage, including recommendations about strategies for prescribing antibiotics in clinical practice, such as the selection of adequate antibiotics, and the dosage or duration of the therapy (3, 4). The ASP-teams, which are composed of specialists from various clinical fields, like clinicians, pharmacists, hygienists, and managers, are considered fundamental to achieve optimization in the rational use of antibiotics within hospitals (5). A growing amount of evidence shows that ASPs can

both optimize the management of infections and reduce the emergence of antimicrobial resistance (6). Nevertheless, a well-known disadvantage of ASPs is the amount of time required to review and document antibiotic therapy (7). There is therefore an urgent need to apply more effective tools to better support clinicians in the complex task of choosing the most appropriate antibiotic treatment. Machine learning techniques and the increasing availability of high-quality, large-scale data offer new opportunities for optimizing antibiotic therapy (8). Thus, decision support systems (DSSs) based on artificial intelligence (AI) use largely automated general learning procedures to identify statistical regularities from the (training) data presented to them and in turn generate predictive probability statements for the occurrence of phenomena (9). The use of AI-based DSSs can be a key factor in improving the results of ASP-teams, considering the multi-user perspective of the problem, the need for knowledge integration from different sources, and the requirement to provide support both for a particular patient and for the whole institution in a coordinated manner. Despite the promising evidence, studies show that there remains some level of inconsistency about the relative merits of AI-based DSSs in influencing practice patterns in hospitals, how to implement them, and what refinements are needed to tailor the systems to local contexts (10). However, the predominant focus on technical prowess during AI development often sidelines considerations for seamless integration into real-world workflows and the practical value of these innovations. If AI-based DSSs are ever to be integrated successfully, it will be essential to establish suitable conditions and develop an adequate strategy for the implementation.

This Delphi study aimed to determine factors that could influence the implementation of AI-based DSSs for antibiotic prescribing in hospitals from the perspective of experts in the field of antibiotic therapy and AI-based DSSs in practice and research.

2 Methods

A two-round Delphi study was undertaken to establish expert consensus on the importance of factors on DSS implementation for antibiotic prescribing in the hospital setting. The Delphi technique is a research method where sequential surveys are used to gain individual expert opinion across several rounds whereby the anonymized results are provided as feedback to the participants (11, 12). In this sense, from the second round of the survey onwards, the experts make their judgements under the influence of the opinions of the other participating experts. In this respect, the Delphi method represents an iterative procedure in which expert assessments on a specific question are determined with the aim of recording and justifying consensus

---

Abbreviations

AI, artificial intelligence; ASPs, antimicrobial stewardship programs; DSSs, decision support systems; GDPR, general data protection regulation; HOT-fit model, human-organization-technology-fit model; XAI, explainable artificial intelligence.

and/or dissent in the judgements (13). The benefits of this method include the possibility of gaining the perspective of an expert group and build consensus in an area where evidence may be lacking (14). The process of the Delphi is illustrated in the following (Figure 1). Ethics approval to conduct the study was obtained from the University of Siegen, project number: LS_ER_25_2024.

2.1 The Delphi instrument

A modified approach was used for this study. The traditional qualitative approach for the first round was discarded (15). Instead, factors for examination were conducted based on two previous studies. The first study involved a systematic review of all literature where implementation factors of AI-based DSSs for antibiotic prescribing in hospitals were described (16). In the second study, a qualitative study of clinicians' perspectives on implementation factors was conducted (17). The preliminary work formed the Delphi instrument used in round one and to provide arguments relating to the factors that could facilitate or hinder the implementation of AI-based DSSs for antibiotic prescribing in hospitals and to which the level of agreement by the participants was acquired. The arguments were grouped into the three domains of the Human-Organization-Technology-(HOT)-fit-model (18), namely human, organization, and technology. Responses were collected with a five-level Likert scale for the evaluation of each thesis and argument: "Disagree" (category response value=1), "rather disagree" (= 2), "rather agree" (= 3), "agree" (= 4) and "I cannot assess" (= 5). An open text box for additional comments related to the topic was provided requiring further investigation in the following round. In the second round, experts were asked to assess the aggregated results of the first round and additionally rank the arguments about the importance for implementation as well as assess the level of realization in current clinical practice and to provide possible interventions for the improvement of current realization. The Delphi instrument was pre-tested for comprehensibility and time required for responses by two people of the target group who did not participate in the main survey. The web-based survey, which was conducted through LimeSurvey was initiated in July and closed in October 2024.

2.2 Survey participants and recruitment

The panel was arranged using purposive sampling (19). According to this, the recruitment focused on expertise and diversity of perspectives rather than representativeness and large sample size (20). So, a heterogeneous panel of experts from research and practice has been compiled. For the study, experts were defined as people who are in a position to assess developments in the healthcare sector concerning the use of AI-based DSS for antibiotic therapy in hospitals. The expert status of the practice-related experts was defined by a professional position or clinical expertise in the field of antibiotic therapy
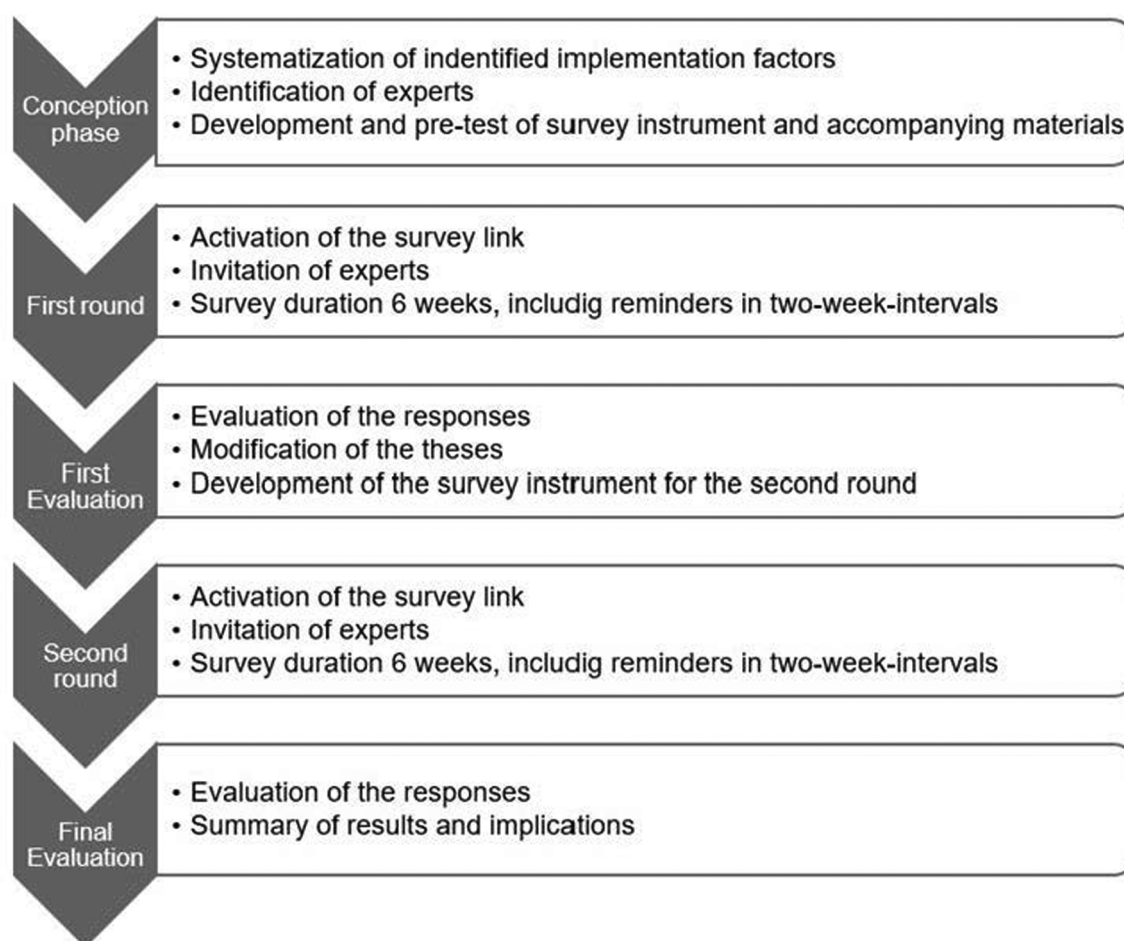
FIGURE 1
The Delphi process.

(e.g., by working as an ASP-member or corresponding professional responsibility in the field of antibiotic therapy). The expert status of scientific experts is defined by scientific achievements, e.g., through publications, special lectures, third-party funded projects, or a proven field of research related to AI-based DSS in healthcare and/or antibiotic therapy e.g., through position title/affiliation that reflects involvement). Specific knowledge of AI-based DSSs was not required, since the term "AI-based DSS" has been explained in the invitation to the survey. 511 experts were invited from across Germany to take part in the study and invitation emails were sent to the individually identified potential participants. Additionally, reminder emails were sent in two-week intervals to all contacts to incentivize those who have not yet participated to attend the survey.

2.3 Consensus and statistical analysis

An analysis of responses was performed after each survey round. As there is no universally defined level of agreement for consensus (19) and a systematic review including 100 Delphi studies found that percent agreement was most frequently used (21), a consensus has been determined according to the rating scale proposed by Meskell et al. (22). The proposed scale is based on preliminary work by de Loë (23), which is also recommended in several studies (24, 25). Consensus is achieved by initially calculating the percentage agreement on items by adding up the case number that achieved the same rating and calculating the percentage. Thus, the level of consensus can be "high", "moderate" or "low" and the direction of consensus can be "in favor" [+] or "against" [−]. A consensus on an argument or thesis is deemed to have been achieved when the group responses reach the predefined percentage threshold. For example, if more than 70% of participants rate an item in one category (e.g., "strongly agree") then the item has a high consensus level. Similarly, if less than 50% of participants rate the item in the category "strongly agree" then the item has a low consensus level. If no consensus was reached, the individual agreement values have been reported. In the next round, the factors for which the responses in the first round resulted in dissent or a moderate or weak consensus (±2% distribution around cut-off value) are submitted for renewed assessment. In

addition, the first-round analysis included a thematic analysis of the open-text-comments to identify additional aspects for inclusion in the second round or reformulating the arguments. The responses of the participants were analyzed, who fully completed the survey. The results were evaluated using SPSS version 24.0 and descriptive statistics. Group differences in the first round were analyzed between the fields of profession (clinical practice and research) using Mann–Whitney-*U*-tests. The Henry Garret ranking method technique (26) was applied to discover the most important factors influencing the implementation of AI-based DSSs for antibiotic prescription in hospitals.

3 Results

77 of the 511 potential participants (15%) invited to take part in the survey accepted the invitation and participated in the first round. 36 of the 77 experts (47%) who initially completed the first round completed the second round. Participant demographics and professional background are presented in Table 1.

3.1 Potential of AI-based DSSs for antibiotic therapy in hospitals

The experts were asked how they would assess the various potentials of AI-based DSSs for antibiotic prescribing in hospitals (Figure 2).

In the first round, the experts agree about the fundamental and divers potentials of AI-based DSSs for antibiotic prescription in hospitals. Thus, the participants grade the potential of AI-based DSSs for providing guidance in case of uncertainty (88%) and improving work processes (87%) as high or rather high. Furthermore, 82% of those surveyed assess the potential of AI-based DSSs for improving the flow of information between providers and thus increasing treatment safety (84%) as high or rather high. The assessment of the potential related to reducing costs is less consistent. Here, 65% of the respondents assume the potential as being high or rather high, whereas 18% assess this potential as being low or rather low and 17% couldńt give an assessment. In the second round, all participants ($n = 36$) agree with the overall picture of the first-round results.

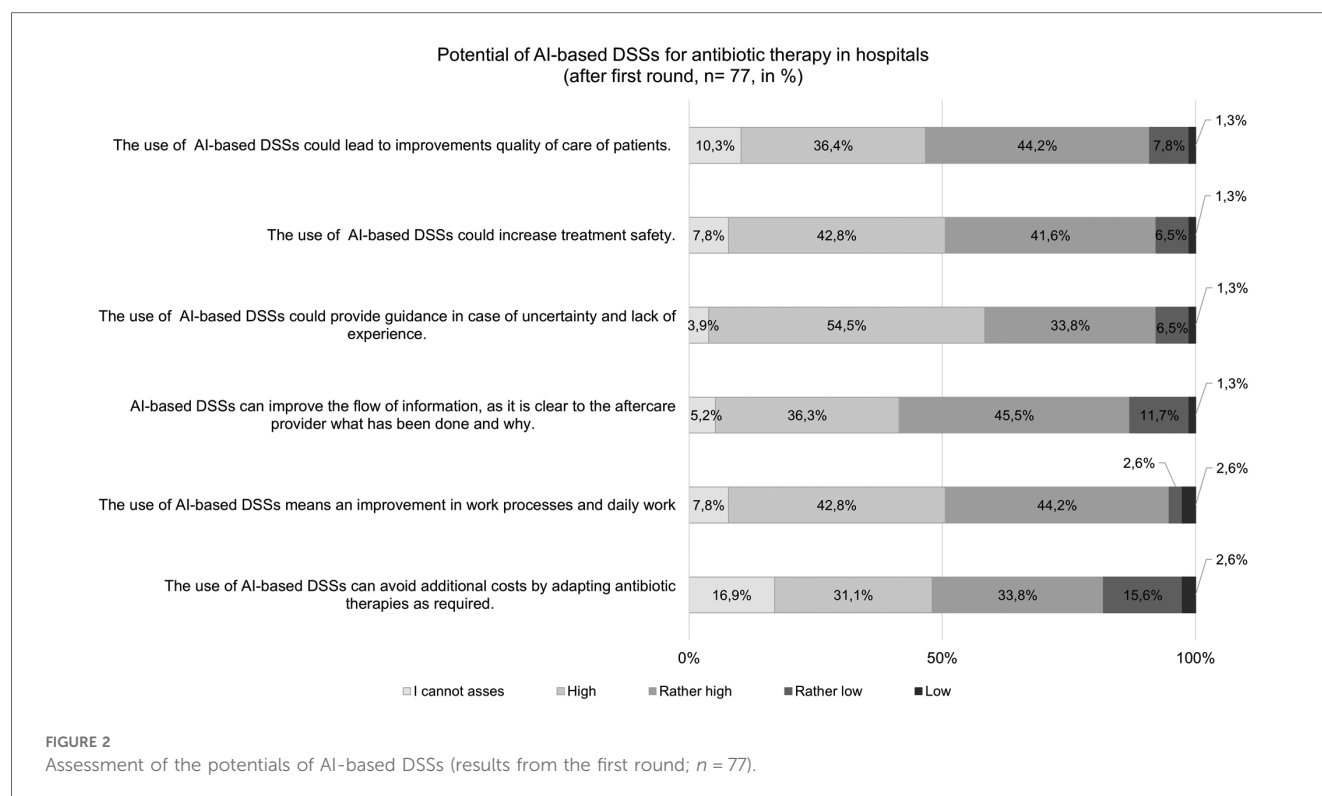## 3.2 Factors of implementation

The following tables show the key results for each of the arguments that could influence AI-based DSS implementation. The arguments are grouped in the domains of the HOT-fit model, and arranged in ascending order by level and direction of consensus.

Those who took part in the Delphi study are uniformly in favor of the aspect that for implementation of AI-based DSSs data of sufficient quality and their regularly maintenance should be given (Table 2).

In addition, they expect AI-based DSSs to be uncomplicated in use and easy to operate as well as recommendations to be

TABLE 1 Participants characteristics.

Demographic		Round 1 ($n = 77$)	Round 2 ($n = 36$)
Gender	Female	36.4% ($n = 28$)	33.3% ($n = 12$)
	Male	63.6% ($n = 49$)	66.7% ($n = 24$)
Age	20–29 years	1.3% ($n = 1$)	2.8% ($n = 1$)
	30–39 years	19.5% ($n = 15$)	25.0% ($n = 9$)
	40–49 years	35.1% ($n = 27$)	25.0% ($n = 9$)
	50–59 years	32.5% ($n = 25$)	38.9% ($n = 14$)
	60 years or older	11.7% ($n = 9$)	8.3% ($n = 3$)
Current professional role	Hospital/Health care services (e.g., an ASP-member)	77.9% ($n = 60$)	75.0% ($n = 27$)
	Technology-related hospital services (e.g., staff in the hospital IT)	2.6% ($n = 2$)	–
	Academic/researcher (e.g., in the fields of AI-based decision support systems)	16.9% ($n = 13$)	22.2% ($n = 8$)
	Other (e.g., hospital hygiene)	2.6% ($n = 2$)	2.8% ($n = 1$)
Years of experience	Up to 5 years	13.0% ($n = 10$)	11.1% ($n = 4$)
	Up to 10 years	26.0% ($n = 20$)	27.8% ($n = 10$)
	More than 10 years	61.0% ($n = 47$)	61.1% ($n = 22$)
Level of expertise in the field of antibiotic therapy	Very high	36.4% ($n = 28$)	44.5% ($n = 16$)
	High	42.9% ($n = 33$)	33.3% ($n = 12$)
	Moderate	9.1% ($n = 7$)	5.6% ($n = 2$)
	Low	5.2% ($n = 4$)	8.3% ($n = 3$)
	Non-specialist	6.5% ($n = 5$)	8.3% ($n = 3$)
Level of expertise in the field of AI-based DSSs	Very high	6.5% ($n = 5$)	2.8% ($n = 1$)
	High	10.4% ($n = 8$)	13.9% ($n = 5$)
	Moderate	36.4% ($n = 28$)	44.5% ($n = 16$)
	Low	35.1% ($n = 27$)	30.5% ($n = 11$)
	Non-specialist	11.7% ($n = 9$)	8.3% ($n = 3$)

FIGURE 2
Assessment of the potentials of AI-based DSSs (results from the first round; *n* = 77).

presented clearly and comprehensible. Although expert consensus demonstrates the importance of patient-specific recommendations, several participants highlight the possible conflicts of DSSs to make specific recommendations and at the same time being too difficult to operate with:

> "The system would have to be far too complicated and confusing for truly individualized recommendations." (Participant 19)

The experts are confident that political and legal frameworks are required for establishing AI-based DSSs. They also believe that supporting users in technical matters and maintaining the system is essential for long-term implementation (Table 3).

Moreover, to increase skills in handling AI-based DSSs, sufficient training opportunities and ongoing familiarization should be offered for (potential) users. Beneath that, AI-based DSSs should be integrated in such a way to support established workflows. Therefore, collateral systematic organizational development and overcoming hierarchical structures are required. The respondents agree likewise that the willingness of hospitals to integrate new technologies into healthcare and the support from clinical leaders can facilitate the implementation. There is no consensus on whether financial incentives for hospitals on the decision to implement and use AI-based DSSs could serve as facilitators. In this context, 30.6% of the participants rather disagree and 8.3% disagree, that monetary incentives are essential for the successful implementation of AI-based DSSs in hospitals, whereas 19.4% agree and 38.9% of the participants rather agree.

The experts believe that AI-based DSSs can be implemented successfully in the long term if users perceive an added value of the use and benefit noticeably from it. Promoting skills in operating with AI-based DSSs, openness towards its use, and trust in it are found to favor implementation (Table 4).

Though there is consensus on previous experience with AI-based DSSs to be beneficial for implementation, it must be considered how users perceive the experience:

> "Previous experience with AI-based DSSs can hinder implementation if it [experience] was negative." (Participant 2)

However, the experts remain in disagreement as to what extent the long-standing professional experience of (potential) users could be a hinderance to the implementation. Here 38.9% of the participants rather disagree and 8.3% disagree with the statement, that the more professional experience users have, the greater their skepticism towards AI-based DSSs might be, which can have a negative impact on successful implementation, while 19.4% of the participants agree and 30.6% rather agree.

3.3 Differences between clinical practice and research

The comparison of the professional role of clinical practice and research (*n* = 63 vs. *n* = 14) revealed differences in the assessment of the implementation factors (Additional File 1 Tables A1, A2). The mean value difference on the following factors are significant: manageable user interface with easy navigation

TABLE 2 Results for the assessed technology-related theses.

Topics and arguments	Consensus*	Direction
Technological factors		
Clear presentation of the results: AI-based DSSs should be attractively designed, and recommendations should be clearly prepared and presented.	(+++)	In favor
Reliable database: The basis for the beneficial use of AI-based DSSs is data of sufficient quality and its regular maintenance.	(+++)	In favor
System integration compatibility: AI-based DSSs should be based on solid data models, interoperable formats, and internationally recognized coding standards so that they can be integrated into any computer system.	(+++)	In favor
Warning functions: To provide helpful support in everyday working life, warning functions of AI-based DSSs should be specific.	(+++)	In favor
Existence of alternative suggestions: An equivalent treatment alternative should be listed to be able to act, e.g., in the case of intolerances and allergies.	(+++)	In favor
Data security: Data protection and security should be regulated in a detailed and generally understandable manner.	(+++)	In favor
Precise recommendations: With AI-based DSSs precise recommendations should be made and direct instructions given on what exactly to do.	(+++)	In favor
Manageable user interface with easy navigation: AI-based DSSs must be uncomplicated to use and easy to operate.	(+++)	In favor
Rapid system updating: AI-based DSSs should be able to be updated quickly.	(+++)	In favor
Automated data transfer: AI-based DSSs are most effective and beneficial when their data is transferred automatically from existing hospital information systems.	(+++)	In favor
Easy access to the system data: AI-based DSSs should be easily accessible and approachable from all common end devices such as smartphones/tablets.	(+++)	In favor
Traceability of recommendations: Users should be provided with comprehensible and expert-curated decision support for possible therapeutic options.	(+++)	In favor
Completeness of recommendations: The recommendations should be comprehensive and complete (e.g., dosage recommendation, duration of therapy).	(+++)	In favor
Easy manual data entry: It is important that AI-based DSSs are designed in such a way that no programmer is required to insert or modify data.	(++)	In favor
Individual-specific recommendations: AI-based DSSs should offer recommendations for specific or individual patient problems.	(+)	In favor

Threshold for consensus in accordance with Meskell et al. (21):
Consensus*:
"High" (+++) >70% in category 4 or 3; >80% in categories 4 and 3.
"Moderate" (++) >60% in category 4 or 3; >70% in categories 4 and 3.
"Low" (+) >50% in category 4 or 3; >60% in categories 4 and 3.
"High" (—) >70% in category 2 or 1; >80% in categories 2 and 1.
"Moderate" (–) >60% in category 2 or 1; >70% in categories 2 and 1.
"Low" (-) >50% in category 2 or 1; >60% in categories 2 and 1.
"Dissent" (x): No unanimous group response.
$N_1 = 77$; $N_2 = 36$.
[a]Results of the second round.

(rather agree: 98.4% clinical practice vs. 78.6% research; $p = 0.003$); completeness of recommendations (agree: 76.2% clinical practice vs. 35.7% research; $p = 0.021$); assurance of technical support (agree: 87.3% clinical practice vs. 64.3% research; $p = 0.018$); promotion of the openness of potential users (agree: 74.6% clinical practice vs. 42.9% research; $p = 0.018$). The differences of the remaining implementation factors are not significant.

3.4 Ranking of implementation factors

To find the most significant factors influencing the implementation of AI-based DSSs for antibiotic prescribing in hospitals, Garrett's Ranking Technique is employed. To do this, the estimated percentage value was converted into point values using the Garrett table. The percentage score is calculated under the following formula (26):

$$\text{Percentage Score} = \frac{(R_{ij}\text{-}0,\,5)}{N_j}$$

Where, R_{ij} = Rank given for i^{th} item j^{th} individual
N_j = Number of items ranked by j^{th} individual.

It is calculated as a percentage score (0–100) and the scale value is obtained by employing the Scale Conversion Table given by Henry Garrett (26). The higher the average score, the more important the factor under consideration (26). It is clear from Figure 3, that in the domain of technology the participants have given importance to the factors (the mean score is indicated in the brackets) system access (78.22), user interface (76.72), and system integration in existing technical structures (74.69).

Figure 4 shows the ranking results of organizational implementation factors that have reached consensus after the first round of the survey. Ensuring technical equipment is occupied the rank 1 (80.92), training offers for (potential) users occupied the rank 2 (74.58) and hospitals̒ willingness to change and integrate new technologies occupied the rank 3 (74.31). Restructuring of medical education (64.92) occupied the lowest rank in the table.

With a view to the ranking of user-related factors that have reached consensus after the first round of the survey, it is evident that the perceived added value of AI-based DSSs (82.14) is the most important implementation factor, followed by trust in its functioning (80.89) and the openness of users to operate with it (77.83). With an average Garret score of 71.08 previous experience with AI-based DSSs is the least (Figure 5).

TABLE 3 Results for the assessed organization-related theses.

Topics and arguments	Consensus*	Direction
Organizational factors		
Hospital's willingness to change: The willingness of hospitals (or of the decision-makers in hospitals) to integrate new technologies into healthcare is essential for AI-based DSS implementation.	(+++)	In favor
Restructuring of "traditional" working routines: AI-based DSSs should be integrated in such a way to support established workflows. A collateral systematic organizational development is therefore required.	(+++)	In favor
Openness of (medical) teams/units: The openness of (medical) teams within an organization towards AI-based DSSs can promote its implementation.	(+++)	In favor
Technical equipment: Sufficient equipment, e.g., in the form of computers/laptops and PC workstations, is a favorable aspect for the implementation of AI-based DSSs.	(+++)	In favor
Availability of technical support: Supporting users in technical matters and the maintenance of the system is essential for implementation.	(+++)	In favor
Clarification of the legal framework: Regulatory control (regulation) and clarification of liability issues are essential for long-term implementation.	(+++)	In favor
Restructuring medical education: The holistic development of AI skills requires further and advanced training programs as well as focal points in medical education.	(+++)	In favor
Support from the management level: Support from clinical leadership/hospital board representatives can facilitate the implementation of AI-based DSSs.	(+++)	In favor
User participation in the development and implementation phase: Future users must be involved in the design and implementation of AI-based DSSs.	(+++)	In favor
Training of potential users: To increase skills in dealing with AI-based DSSs, sufficient training opportunities should be offered for users as well as ongoing familiarization.	(+++)	In favor
Overcoming hierarchical structures: Hierarchical structures and dependence on established standards are an obstacle to the implementation of AI-based DSSs.	(+)	In favor
Financial incentives: Increased monetary incentives for the implementation and use of AI-based DSSs in hospitals are essential for successful implementation.[a]	(x)	

Threshold for consensus in accordance with Meskell et al. (21):

Consensus*:

"High" (+++) >70% in category 4 or 3; >80% in categories 4 and 3.

"Moderate" (++) >60% in category 4 or 3; >70% in categories 4 and 3.

"Low" (+) >50% in category 4 or 3; >60% in categories 4 and 3.

"High" (—) >70% in category 2 or 1; >80% in categories 2 and 1.

"Moderate" (–) >60% in category 2 or 1; >70% in categories 2 and 1.

"Low" (-) >50% in category 2 or 1; >60% in categories 2 and 1.

"Dissent" (x): No unanimous group response.

$N_1 = 77$; $N_2 = 36$.

[a]Results of the second round.

TABLE 4 Results for the assessed user-related theses.

Topics and arguments	Consensus*	Direction
User-related factors		
Promoting of competencies in operating with AI-based DSSs: The degree of perceived competencies among users influences the successful implementation of AI-based DSSs.	(+++)	In favor
Reduction of uncertainties: Due to reservations about AI-based DSSs, assessments could be categorically questioned by users, which can lead to DSSs not being used effectively.	(+++)	In favor
Openness of potential users: The openness of potential users towards AI-based DSSs is essential for implementation in hospitals.	(+++)	In favor
Knowledge and understanding of how AI-based systems work: It is important that human expertise is paired with the recommendations of AI-based DSSs (i.e., qualified people validate the results).	(+++)	In favor
Perceived added value of the use of AI-based DSSs: Only if users benefit noticeably from AI-based DSSs, can be implemented successfully in the long term.	(+++)	In favor
Trust in the functioning of AI-based DSSs: Trust of users in AI-based DSSs and its recommendations are essential for implementation.	(+++)	In favor
Previous experience with AI-based DSSs: Previous experience with AI-based DSSs is beneficial for successful implementation.	(++)	In favor
Age of users: The age of users influences the implementation and use of AI-based DSSs.[a]	(++)	In favor
Professional experience: The more professional experience users have, the greater the skepticism towards AI-based DSSs, which can have a negative impact on successful implementation.[a]	(x)	

Threshold for consensus in accordance with Meskell et al. (21):.

Consensus*:

"High" (+++) > >70% in category 4 or 3; >80% in categories 4 and 3.

"Moderate" (++) >60% in category 4 or 3; >70% in categories 4 and 3.

"Low" (+) >50% in category 4 or 3; >60% in categories 4 and 3.

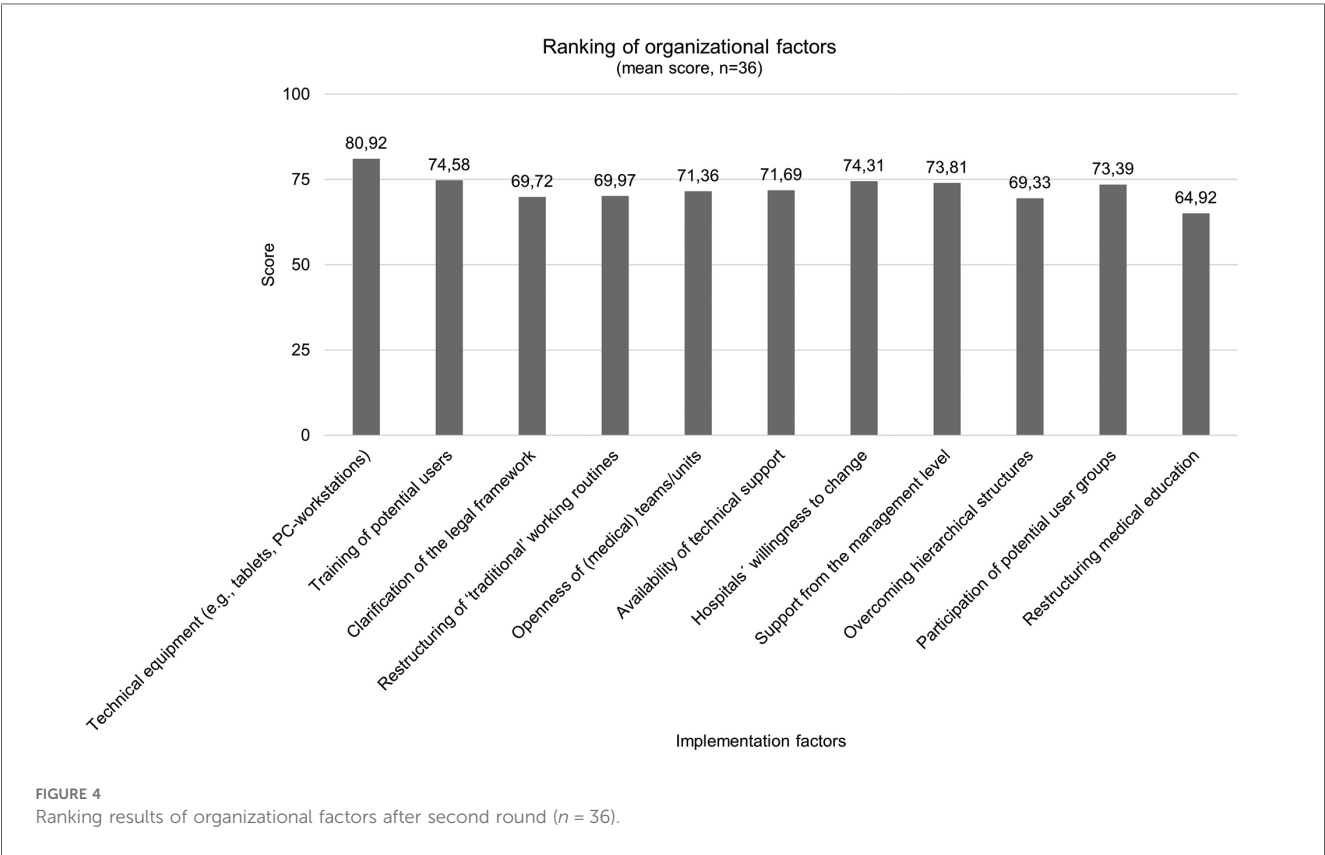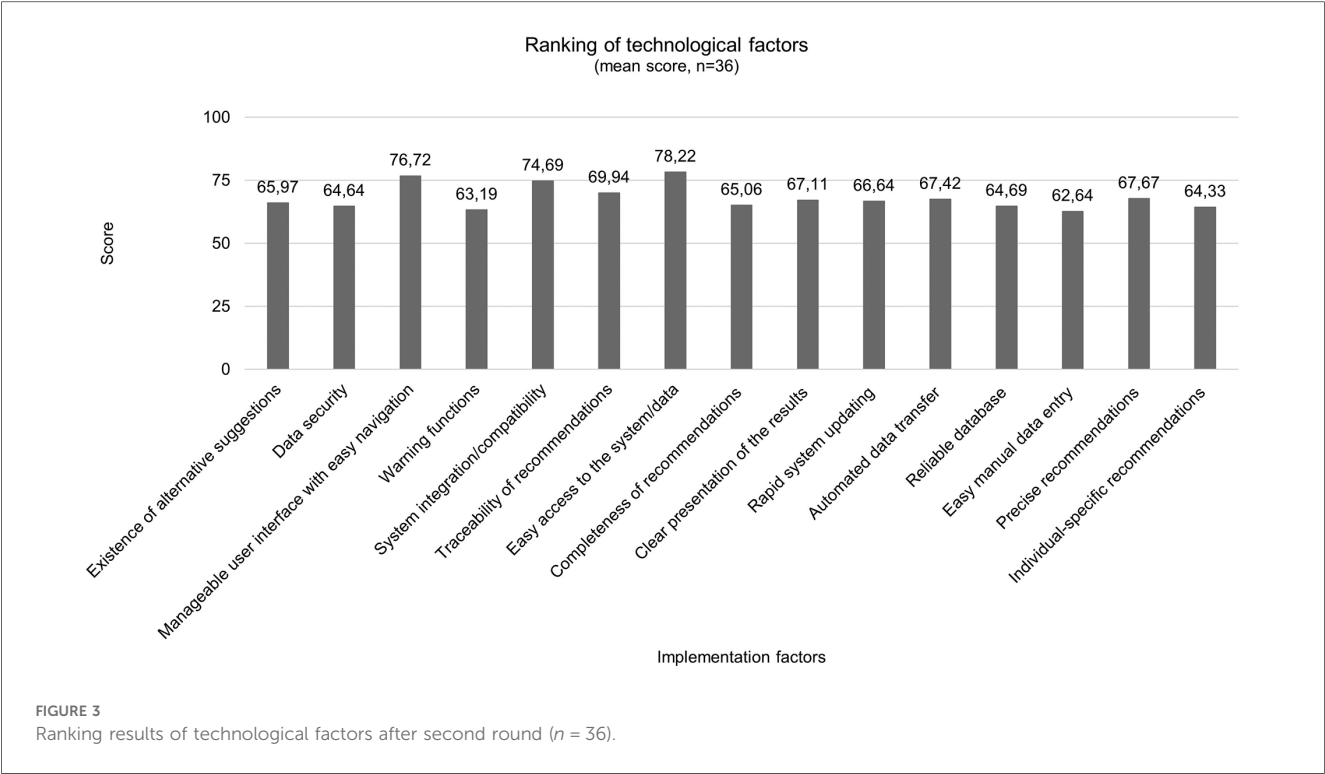"High" (—) >70% in category 2 or 1; >80% in categories 2 and 1.

"Moderate" (–) >60% in category 2 or 1; >70% in categories 2 and 1.

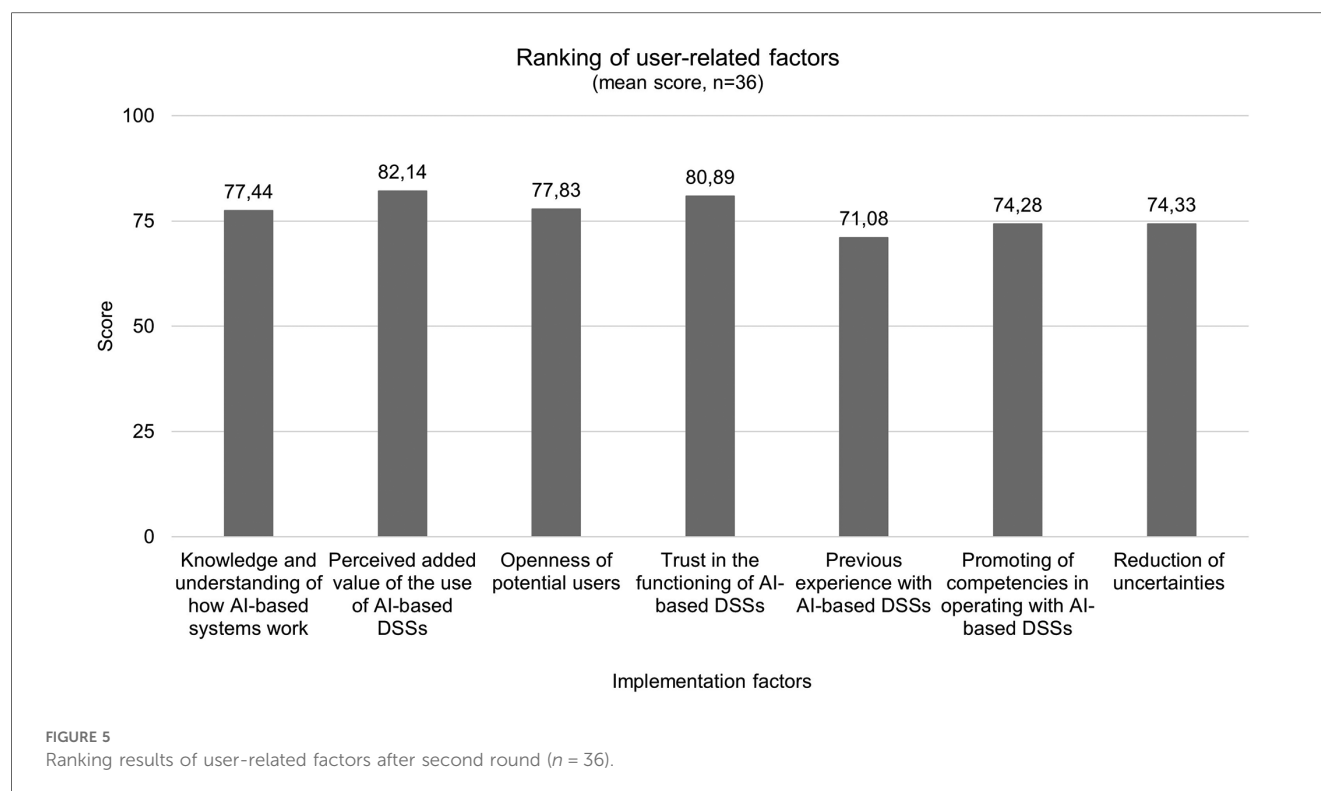"Low" (-) >50% in category 2 or 1; >60% in categories 2 and 1.

"Dissent" (x): No unanimous group response.

$N_1 = 77$; $N_2 = 36$.

[a]Results of the second round.

FIGURE 3
Ranking results of technological factors after second round (*n* = 36).

FIGURE 4
Ranking results of organizational factors after second round (*n* = 36).

FIGURE 5
Ranking results of user-related factors after second round (*n* = 36).

The details of the ranking of factors influencing the implementation of AI-based DSSs frequencies of the Henry Garret Ranking method are obtained in supplementary material (Additional File 2 Tables A1–A10).

3.5 Current realization of factors

The respondents were then asked for their assessment of the current realization of the factors that reached consensus in the first round. The results show that the participants mostly were unable to give a clear assessment of implementation. For example, eight respondents consider the factor that hospitals are supported in their willingness to change to be rather well realized, while six respondents rate this factor as not yet realized at all, and 15 respondents are unable to assess in this regard. In addition, ten participants rate ensuring data security as very good or rather good, while three participants rate it as rather not well realized and 23 could give no appraisal. Moreover, 13 respondents assess the sharing of knowledge in the context of AI-based DSSs as rather poorly or poorly realized, while nine respondents consider this factor as not realized at all, and 12 respondents could not give an assessment. The details of the appraisal of the current implementation of factors can be gathered in Table A1 in Additional File 3.

When asked with which interventions and strategies implementation might be promoted, one participant states, that developers should offer open interfaces that are freely available to every institution, instead of every development team having and providing their own interfaces, licenses, and databases, that cannot

be acquired and used. Besides, the majority of AI-based DSSs that are currently being developed as prototypes by researchers are not comprehensible. More research in the field of explainable AI (XAI) could be a favorable factor in terms of traceability. In this context, in the development phase revision loops with a strong emphasis on usability evaluation should be pursued as well as the definition and effective application of interoperability standards. Moreover, the definition of a legal framework would be lagging behind the technical development, so it appears to be important for the state to be more active and to invest in the development of a cornerstone for the implementation and use of AI-based DSSs-meaning the technical and legal infrastructure and also to invest in the improvement of the digitalization level of hospitals, that are not operating with such systems, yet. In addition, the adaptation of the medical training should be envisaged as well as the establishment of multidisciplinary teams that would have been mandated with the support and monitoring of the implementation. Finally, raising awareness among (potential) users should be taken seriously to break the distorted facts about the use of AI and to report objectively on the possibilities and limitations of its use. This could be realized through e.g., workshops or technology showrooms.

3.6 Factors impacting trust in AI-based DSSs

Respondents to the current survey share their views about how to build trust in AI-based DSSs (Figure 6). Almost half of the respondents say training AI-based DSSs to be factually accurate, moral, and not harmful would rather increase and more than a
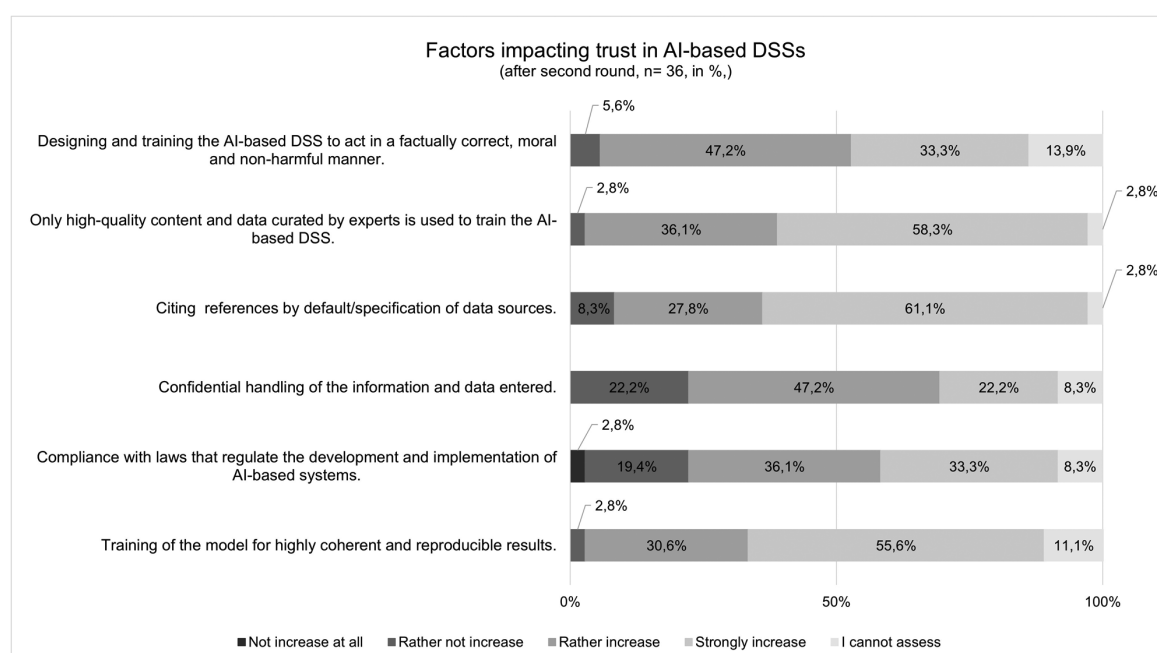
FIGURE 6
Assessment of factors influencing trust in AI-based DSSs (*n* = 36).

third of the respondents say would strongly increase their trust in that system. Additionally, almost 60% of respondents say only using high-quality content to train DSSs would strongly increase their trust, while 55% say training the model for highly coherent and reproducible results would strongly increase their trust. Transparency is also an important factor so for more than 60% of the respondents, citing references by default will strongly increase trust. The confidential handling of the information is rather a trust-facilitating factor for 47% and strongly facilitating for 22% of the respondents. More than a third of the respondents assessed compliance with laws as strongly increasing their trust.

In conclusion, looking to the future, 75% of the respondents (*n* = 27) agree that AI-based DSSs for antibiotic prescribing in hospitals will gain significance, while 25% consider them to gain slight importance.

4 Discussion

This study aimed to systematically evaluate AI-based DSSs and the factors that could potentially influence their implementation for antibiotic prescribing in hospitals from the perspective of a Delphi panel. The experts recognize the increasing importance of corresponding systems. The participants expect AI-based DSSs to be more present and incorporated into daily work routines in the future. The theses presented in the study can be combined to paint an overall picture of how AI-based DSSs might be incorporated into care and research in the future. The participants assess many of the arguments relating to the potential of AI-based DSSs positively. The experts see an added value in AI-based DSSs when it comes to providing guidance for antibiotic prescription as well as increasing treatment safety and improving daily work. Nevertheless, it has to be taken into account that adopting a supposedly correct result without proper consideration might provide the potential for medication errors (27). Even though DSSs facilitate decision-making and have the potential to reduce the workload, the users' ability to act may be limited, when the system fails and the degree of reliance on the recommendations can generate a form of dependency as well (28). It is clear that the decision-making process of clinicians cannot be completely replaced by AI-based DSSs, nor that this would be desired by those involved (17, 29). Here it is necessary to clarify the tasks of AI-based DSSs in a participatory process and enable a continuous process of implementation.

Consensus has been established for 34 implementation factors: 15 technological factors, 11 organizational, and eight user-related factors. The respondents have then given their preferential ordering of the unranked implementation factors that reached a consensus at the initial stage. Garret´s Ranking Technique revealed that the most important factor influencing implementation in the domain of organization was the assurance of technical equipment with a mean score of 80.92. This also contains hospitalś willingness to incorporate innovations, which reached rank three in this study, and also to integrate responsible authorities, like the management level in this process to establish organizational structures for user-friendly implementation (30). When presented with the issue of "clarification of the legal framework", the panel is in agreement to achieve this objective. The progress of AI in healthcare has consequences for the law. The legal challenges are also based on the risks associated with

the systems, such as unpredictability and uncontrollability (31). From a legal perspective, the diffusion of responsibility, the relationship between healthcare professionals and patients as well as regulations concerning personal data are relevant topics (32). The protection of personal data is regulated in detail by the General Data Protection Regulation (GDPR) (33). Nevertheless, there is still a need for further discussion in context of AI-based systems. For example, the right to erasure under Art. 17 GDPR poses considerable difficulties, since the complete deletion of personal data that has already been processed from an AI-based system is technically difficult due to its technical architecture (34). An adaptation to the architecture of AI should therefore continue to be discussed. Concepts such as dynamic consent, the introduction of data trustees and the possibility of data donation enable the people to have more control over their data (35). Another aspect that is being discussed concerns the allocation of data as a type of special property. In this context, the question would arise as the system may use the data for further learning or whether it may be given to third parties in anonymized form, e. g., for further development of the system (36). Consequently, fundamental changes will be required, since there is also a lack of clarity concerning what authority and responsibility the users must assume. There are uncertainties regarding the distribution of roles and responsibilities that might be intensified by the use of AI-based DSSs, especially dealing with the recommendations as well as their binding nature, which possibly reduce the acceptance of the system (37). To address this, solutions for an appropriate distribution of responsibility are needed, whereby all relevant parties like developers, clinicians as (potential) users, hospitals, and their managers and experts for data protection and ethics (33). More attention has also to be paid to further education and training of medical staff in particular dealing with AI-based systems and the provision of necessary information. Linked with that, there is consensus among the experts that knowledge and competencies will be key factors in the future. A consensus study of a working group on AI in healthcare also concluded, that educators in healthcare should define new competencies for using AI-based DSS and these requirements should be incorporated into medical education (38). In this context, it is about digital literacy, e.g., the competent handling of the system and execution of the recommendations in practice, which has so far also lacking in medical education (39), even if this factor has been ranked the least important among the experts of this Delphi study. Moreover, the factor "perceived added value of AI-based DSSs", has been ranked as most important in the user-related domain. In the context of evidence-based medicine, the safety, appropriateness, and efficacy of new interventions must be proven before they can be used in clinical practice. Following this paradigm, AI-based DSSs should also have an additional benefit before they can become part of routine care and be fully accepted by clinicians. Additionally, userś trust in DSSs has been ranked as the second important factor influencing implementation, which is in line with previous studies (40, 41). Trust can be achieved by creating transparency regarding the extent to which the recommendations of DSSs apply to the context in which they are embedded in a value-oriented manner (42). Explainability is but one indicator of

achieving transparency as well as an open communication on the systemś capabilities and limitations (43). Methods from the field of XAI are important here. The so-called global explanation describes which variables are particularly relevant for the model in general and how strongly they influence the systemś recommendations (44). However, a specific decision can also be based partly on other factors - this is referred to as local explanation (45). Although neither of these factors makes it possible to open the "black box", together with information on the training data, the methodology used and the results of validation studies, it increases the transparency of the system and gives users an indication of how it works internally, which can promote trust in DSSs (46). In this issue, a study analyzing the development and implementation of a DSS for antimicrobial stewardship in two Swiss hospitals highlights the fact, that making the underlying process of decision-making transparent and understandable for clinicians they were more willing to accept the system (47). In addition, it was beneficial to keep the process simple and visible for users as well as make clear for them where the recommendations, respective the content is coming from (47). Not all users need to understand the system as a whole to trust it, but they should know why they can trust it (48). This is also reflected in the respondents' assessments of factors that influence trust in DSSs. Here, the data and evidence on which the recommendations are based play significant roles for the majority of respondents. However, evidence-based medicine does not only refer to the effectiveness of the respective systems as determined by scientific standards but also to linking this to the expertise of medical staff in the best possible way (49). Concerning technological factors "easy access to the system/data" reached the first rank with a mean score of 78.22. This is in line with the results of another Delphi study, where consensus was sought on a core list of important safety features to be considered when designing, implementing or using DSSs (50). The panelists agreed on items related to the ease of using the system and also that the system should provide prompts to specifying doses of the medications prescribed, frequencies and dosage forms (50), which also reached consensus in this Delphi study.

At this point, it seems to be important to figure out what is technically possible and how the technical possibilities can be utilized most profitably. Generally speaking, DSSs should be designed along with clinical needs and not only according to the developerś possibilities. The fit of the technology with clinical processes was also addressed in a study in which the integration of a DSS in an Austrian hospital was described (51). Here it proved to be a challenge that those developing the technology were not familiar with the organizational processes. Therefore, implementation failed and could only be continued after appropriate adjustments. It also became apparent that the lack of cooperation with clinical stakeholders meant that the added value of DSS for improving quality of care and working routines was not recognized or accepted (51). In addition, it appears that system attributes of AI-based DSSs have to simultaneously meet the demands of e.g., easy access, compatibility, transparency, or explainability, whereby not everything can be optimized at the same time. One opportunity to apply to this demands is to

provide key facts for interested and informed users in the form of an easy-to-understand manual, so-called model cards, about the underlying model, characteristics of the training data, and system evaluation (52). Furthermore, the requirement to make the recommendations of an AI-based DSS more explainable will make it necessary to involve medical staff more in the development of such systems (53), which also reached a consensus within the Delphi panel. It is also important to consider the items that did not reach consensus. In this context, there is no clear tendency relating professional experience as being a hinderance or facilitator for DSS implementation. It might be possible that the sample composition had an influence on the results. In this study, people participated, who are clinical experts and just in the beginning of their professional career and probably more open towards innovative ways of care, as people who have more professional experience and are more reserved towards change. It is well known that in deep-rooted systems and hierarchies characterized by seniority like in hospitals, there might be resistance towards scientific evidence (54). Besides, expertise have also been found to influence perceptions and use of DSS implementation (55), so that the heterogeneity of expert levels in this study have also affected the assessment of implementation factors.

However, as it is apparent from the appraisal of the current realization of the factors if AI-based DSSs for antibiotic prescribing are to be considered reliable, safe, and beneficial to human well-being, comprehensive strategies and frameworks for their use in the healthcare sector will need to be developed over the next few years in collaboration with various stakeholders.

5 Strengths and limitations

This study was the first to systematically determine which factors should be considered when implementing AI-based DSSs in hospital settings from the perspective of experts in the field of antibiotic prescribing as well as AI-based DSSs in practice and research.

Even though the methodology of the Delphi method is an accepted procedure in the scientific community, the results reflect an accumulation of subjective judgements. There is always a degree of uncertainty about current and future developments. The appraisal of influencing factors is based on the expectations of the sample and is not representative (56). The challenging identification and recruitment of experts (especially in the field of development and deployment of AI-based DSSs for the hospital setting) indicates that there are still few specialists in this area. In addition, the survey results largely reflect the assessment of ASP-members in hospitals, as these are overrepresented in the survey. The overrepresentation of clinical experts might have affected the prioritization of implementation factors, insofar that technical requirements were seen as less as a concern than organizational and user-related factors. A more balanced panel would be particularly desirable in the context of the variability of perspectives. Further on, one conclusion from the response rates of 15% in the first and 47% in the second round is that on the one

hand, the two-step process may have been too time-consuming and on the other hand the questionnaire was possibly too extensive. However, this seemed necessary to gain an overview of the manifold factors influencing implementation in healthcare facilities like hospitals. For further research, a shorter and more focused questionnaire should be developed. Still, the panel size was within the range of sizes used in previous studies involving achieving consensus on issues in healthcare (57, 58). The high level of expertise of the sample concerning context-related antibiotic prescription speaks in favor of high informative value. The rather moderate to low expertise in the field of DSSs might be a possible reason for the inconclusive assessment of current realization of implementation factors, since existing knowledge gaps as well as insufficient implementation strategies and activities in German hospitals may have had an impact on the assessment, which is limiting. Nonetheless, the results provide a broad overview from a context-related perspective about what seems to be important and relevant in terms of DSS implementation in hospital setting. That means that the results of the survey can be used to develop tools and frameworks that support the planning and implementation of AI-based DSSs in hospitals and the people involved.

6 Conclusion

With the aid of the Delphi method, it was possible to provide insights into factors related to the implementation of AI-based DSSs for antibiotic prescribing in German hospitals from the perspective of experts, whereby a wide range of factors seem to be crucial. Although there are promising examples of good practices for the use of AI-based DSSs in healthcare available, the extent of professional role-related or even overall societal effects are hardly foreseeable. Requirements of AI-based DSSs in everyday clinical practice are to promote application-orientated development as well as integrate them into clinical organizational structures and fulfill demands of trust, transparency, and responsibility. Given considerable uncertainties regarding the effects of AI-based DSS implementation in healthcare, the organizational, social, and ethical aspects of this transformation process must be given more attention than before, both in practice and research.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by Ethics Committee of the University of Siegen. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PT: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. JA: Conceptualization, Writing – review & editing. CD: Supervision, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2025.1555042/full#supplementary-material

References

1. Rid A, Littmann J, Buyx A. Evaluating the risks of public health programs: rational antibiotic use and antimicrobial resistance. *Bioethics*. (2019) 33(7):734–48. doi: 10.1111/bioe.12653

2. Kern WV. Rationale antibiotikaverordnung in der humanmedizin. *Bundesgesundheitsblat Gesundheitsforschung Gesundheitsschutz*. (2018) 61(5):580–8. doi: 10.1007/s00103-018-2727-x

3. Dyar OJ, Huttner B, Schouten J, Pulcini C, ESCMID Study Group for Antibiotic Policies. What is antimicrobial stewardship? *Clin Microbiol Infect*. (2017) 23:793–8. doi: 10.1016/j.cmi.2017.08.026

4. Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, et al. Implementing an antibiotic stewardship program: guidelines by the Infectious Diseases Society of America and the society for healthcare epidemiology of America. *Clin Infect Dis*. (2016) 62(10):e51–77. doi: 10.1093/cid/ciw118

5. Hulscher MEJL, Prins JM. Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. *Clin Microbiol Infect*. (2017) 23(11):799–805. doi: 10.1016/j.cmi.2017.07.017

6. Hoffmann C, Zak M, Avery L, Brown J. Treatment modalities and antimicrobial stewardship initiatives in the management of intra-abdominal infections. *Antibiotics*. (2016) 5(1):11. doi: 10.3390/antibiotics5010011

7. Palacios F, Campos M, Juarez JM, Cosgrove SE, Avdic E, Cánovas-Segura B, et al. A clinical decision support system for an antimicrobial stewardship program. *HEALTHINF 2016–9th International Conference on Health Informatics (Proceedings)*. Rome: SciTePress (2016). p. 496–501

8. Athey S. The impact of machine learning on economics. In: Agrawal A, Gans J, Goldfarb A, editors. *The Economics of Artificial Intelligence*. Chicago: Harvard Business Press, University of Chicago Press (2018). p. 507–52.

9. Börm P. Leitlinienbasierter clinical decision support—anforderungen an evidenzbasierte entscheidungsunterstützungssysteme. *OP-J*. (2021) 37(1):28–35. doi: 10.1055/a-1284-3193

10. 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. doi: 10.1093/jac/dky222

11. Trevelyan EG, Robinson N. Delphi methodology in health research: how to do it? *Eur J Integr Med*. (2015) 7(4):423–8. doi: 10.1016/j.eujim.2015.07.002

12. Niederberger M, Renn O. Das klassische delphi-verfahren: konzept und vorgehensweise. In: Niederberger M, Renn O, editors. *Das Gruppendelphi-Verfahren*. Wiesbaden: Springer Fachmedien Wiesbaden (2018). p. 7–25.

13. Cuhls K. *Die Delphi-Methode—eine Einführung. Delphi-Verfahren in den Sozial- und Gesundheitswissenschaften, Konzept, Varianten und Anwendungsbeispiele*. Wiesbaden: Springer Fachmedien Wiesbaden (2019). p. 3–31.

14. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. (2011) 6(6):e20476. doi: 10.1371/journal.pone.0020476

15. Keeney S, Hasson F, McKenna H. The delphi technique. In: Keeney S, Hasson F, McKenna H, editors. *The Delphi Technique in Nursing and Health Research*. Oxford: Wiley (2011). p. 1–17.

16. 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):1–12. doi: 10.1186/s12911-023-02124-4

17. Tokgöz P, Hafner J, Dockweiler C. Faktoren für die implementierung von KI-basierten entscheidungsunterstützungssystemen zur antibiotikavorhersage im krankenhaus—eine qualitative analyse aus der perspektive von ärztlichem personal. *Gesundheitswesen*. (2023) 85(12):1220–8. doi: 10.1055/a-2098-3108

18. Yusof MM, Papazafeiropoulou A, Paul RJ, Stergioulas LK. Investigating evaluation frameworks for health information systems. *Int J Med Inf*. (2008) 77(6):377–85. doi: 10.1016/j.ijmedinf.2007.08.004

19. Hasson F, Keeney S, McKenna H. Research guidelines for the delphi survey technique. *J Adv Nurs*. (2000) 32(4):1008–15. doi: 10.1046/j.1365-2648.2000.t01-1-01567.x

20. Steinmüller K. Das "klassische" Delphi. Praktische herausforderungen aus sicht der zukunftsforschung. In: Niederberger M, Renn O, editors. *Delphi-Verfahren in den Sozial- und Gesundheitswissenschaften*. Konzept, Varianten und Anwendungsbeispiele, Wiesbaden, Heidelberg: Springer VS (2019). p. 33–54.

21. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol*. (2014) 67(4):401–9. doi: 10.1016/j.jclinepi.2013.12.002

22. Meskell P, Murphy K, Shaw DG. Insights into the use and complexities of the policy delphi technique. *Nurse Res*. (2014) 21(3):32–9. doi: 10.7748/nr2014.01.21.3.32.e342

23. de Loë RC. Exploring complex policy questions using the policy Delphi: a multi round, interactive survey method. *Appl Geogr*. (1995) 15(1):53–68. doi: 10.1016/0143-6228(95)91062-3

24. O'Loughlin R, Kelly A. Equity in resource allocation in the Irish health service: a policy Delphi study. *Health Policy*. (2004) 67(3):271–80. doi: 10.1016/j.healthpol.2003.07.001

25. Landeta J. Current validity of the Delphi method in social sciences. *Technol Forecast Soc Change*. (2006) 73(5):467–82. doi: 10.1016/j.techfore.2005.09.002

26. Garrett EH, Woodworth RS. *Statistics in Psychology and Education*. Bombay: Vakils, Feffer and Simons Pvt. Ltd. (1969). p. 329.

27. Trivedi MH, Daly EJ, Kern JK, Grannemann BD, Sunderajan P, Claassen CA. Barriers to implementation of a computerized decision support system for depression: an observational report on lessons learned in "real world" clinical settings. *BMC Med Inform Decis Mak*. (2009) 9(6):1–9. doi: 10.1186/1472-6947-9-6

28. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med*. (2020) 3(17):1–10. doi: 10.1038/s41746-020-0221-y

29. Jörg J. *Digitalisierung in der Medizin: Wie Gesundheits-Apps, Telemedizin, Künstliche Intelligenz und Robotik das Gesundheitswesen Revolutionieren*. Berlin Heidelberg: Springer (2018). p. 85–11. doi: 10.1007/978-3-662-57759-2

30. van Gemert-Pijnen JEWC. Implementation of health technology: directions for research and practice. *Front Digit Health*. (2022) 4:1030194. doi: 10.3389/fdgth.2022.1030194

31. Palaniappan K, Lin EYT, Vogel S. Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector. *Healthcare*. (2024) 12(5):562. doi: 10.3390/healthcare12050562

32. Meszaros J, Minari J, Huys I. The future regulation of artificial intelligence systems in healthcare services and medical research in the European union. *Front Genet*. (2022) 4(13):927721. doi: 10.3389/fgene.2022.927721

33. Beck S, Faber M. Rechtlicher rahmen des einsatzes von KI in der medizin. In: Oppermann B, Buck-Heeb P, editors. *Automatisierte Systeme*. München: Beck (2022). p. 267–300.

34. General Data Protection Regulation. Chapter 3, Art. 17 (2025). Available online at: https://gdpr-info.eu/art-17-gdpr/#:~:text=the%20personal%20data%20have%20to, in%20Article%208(1) (accessed March 28, 2025)

35. Hummel P, Braun M, Augsberg S, von Ulmenstein U, Dabrock P. *Datensouveränität: Governance-Ansätze für den Gesundheitsbereich*. Wiesbaden: Springer Fachmedien (2021). doi: 10.1007/978-3-658-33755-1

36. Siems J, Repka T. *Unrechtmäßige Nutzung von KI-Trainingsdaten als Gefahr für Neue Geschäftsmodelle?—compliance-Anforderungen bei der Entwicklung von KI-Systemen. Herbstakademie*. Deutsche Stiftung für Recht und Informatik (2021). p. 517–34. https://dsri.de/storage/siems_repka.pdf (accessed March 28, 2025)

37. Beck S. Künstliche intelligenz—ethische und rechtliche herausforderungen. In: Mainzer K, editor. *Philosophisches Handbuch Künstliche Intelligenz*. Wiesbaden: Springer Fachmedien (2019). p. 1–28.

38. Labkoff S, Oladimeji B, Kannry J, Solomonides A, Leftwich R, Koski E, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. *J Am Med Inform Assoc*. (2024) 31(11):2730–9. doi: 10.1093/jamia/ocae209

39. Stanford Medicine. The Rise of the Data-Driven Physician. (2020). Available online at: https://med.stanford.edu/content/dam/sm/school/documents/Health-Trends-Report/Stanford%20Medicine%20Health%20Trends%20Report%202020.pdf (accessed December 18, 2024)

40. Pope C, Halford S, Turnbull J, Prichard J, Calestani M, May C. Using computer decision support systems in NHS emergency and urgent care: ethnographic study using normalisation process theory. *BMC Health Serv Res*. (2013) 13(1):111. doi: 10.1186/1472-6963-13-111

41. Gretton C. Trust and transparency in machine learning-based clinical decision support. In: Zhou J, Chen F, editors. *Human and Machine Learning. Human–Computer Interaction Series*. Cham: Springer (2018). p. 279–92. doi: 10.1007/978-3-319-90403-0_14

42. von Eschenbach WJ. Transparency and the black box problem: why we do not trust AI. *Philos Technol*. (2021) 34:1607–22. doi: 10.1007/s13347-021-00477-0

43. European Commission. High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI (2019). Available online at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed March 28, 2025)

44. Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S, et al. To explain or not to explain?—artificial intelligence explainability in clinical decision support systems. *PLoS Digit Health*. (2022) 1(2):e0000016. doi: 10.1371/journal.pdig.0000016

45. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*. (2018) 6:52138–60. doi: 10.1109/ACCESS.2018.2870052

46. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain. (2017). doi: 10.48550/arXiv.1712.09923

47. Catho G, Centemero NS, Waldispühl Suter B, Vernaz N, Portela J, Da Silva S, et al. How to develop and implement a computerized decision support system integrated for antimicrobial stewardship? Experiences from two Swiss hospital systems. *Front Digit Health*. (2021) 16(2):583390. doi: 10.3389/fdgth.2020.583390

48. Böhme G. *Invasive Technisierung: Technikphilosophie und Technikkritik*. Kusterdingen: Die Graue Edition (2008).

49. Worf K, Tränkner N, Wilke M. Evidenzbasierte medizin in deutschland—zugangssteuerung oder innovationsbremse für medizinprodukte? In: Pfannstiel MA, Jaeckel R, Da-Cruzi P, editors. *Market Access im Gesundheitswesen: Hürden und Zugangswege zur Gesundheitsversorgung*. Wiesbaden: Springer Fachmedien (2020). p. 135–67.

50. Shawahna R. Merits, features, and desiderata to be considered when developing electronic health records with embedded clinical decision support systems in Palestinian hospitals: a consensus study. *BMC Med Inform Decis Mak*. (2019) 19(1):216. doi: 10.1186/s12911-019-0928-3

51. Schuh C, de Bruin JS, Seeling W. Clinical decision support systems at the Vienna general hospital using arden syntax: design, implementation, and integration. *Artif Intell Med*. (2018) 92:24–33. doi: 10.1016/j.artmed.2015.11.002

52. Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, et al. Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019). p. 220–9. doi: 10.1145/3287560.3287596

53. Fernau S, Schleidgen S, Schickhardt C, Oßa AK, Winkler EC. Zur rolle und verantwortung von Ärzten und forschern in systemmedizinischen kontexten: ergebnisse einer qualitativen interviewstudie. *Ethik Med*. (2018) 30(4):307–24. doi: 10.1007/s00481-018-0494-8

54. Varonen H, Kortteisto T, Kaila M. What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians. *Fam Pract*. (2008) 25(3):162–7. doi: 10.1093/fampra/cmn020

55. Esmaeilzadeh P, Sambasivan M, Kumar N, Nezakati H. Adoption of clinical decision support systems in a developing country: antecedents and outcomes of physician's threat to perceived professional autonomy. *Int J Med Inform*. (2015) 84(8):548–60. doi: 10.1016/j.ijmedinf.2015.03.007

56. Niederberger M, Renn O. Das gruppendelphi- verfahren in den sozial- und gesundheitswissenschaften. In: Niederberger M, Renn O, editors. *Delphi-Verfahren in den Sozial- und Gesundheitswissenschaften. Konzept, Varianten und Anwendungsbeispiele*. Wiesbaden, Heidelberg: Springer VS (2019). p. 83–100.

57. Spyropoulos AC, Viscusi A, Singhal N, Gilleylen J, Kouides P, Howard M, et al. Features of electronic health records necessary for the delivery of optimized anticoagulant therapy: consensus of the HER task force of the New York state anticoagulation coalition. *Ann Pharmacother*. (2015) 49(1):113–24. doi: 10.1177/1060028014555176

58. Shawahna R, Al-Atrash M. What do primary healthcare providers and complementary and alternative medicine practitioners in Palestine need to know about exercise for cancer patients and survivors: a consensual study using the Delphi technique. *Evid Based Complement Alternat Med*. (2019) 2019(3):1–14. doi: 10.1155/2019/7695818