



OPEN ACCESS

EDITED BY

Guglielmo M. Trovato,
European Medical Association (EMA), Belgium

REVIEWED BY

Arijita Banerjee,
Indian Institute of Technology Kharagpur,
India
Cornelius G. Wittal,
Roche Pharma AG, Germany
Dhurandhar,
Pt. Jawahar Lal Nehru Memorial Medical
College, India

*CORRESPONDENCE

Mario Muselli
✉ mario.muselli@univaq.it

RECEIVED 28 June 2025

ACCEPTED 26 August 2025

PUBLISHED 18 September 2025

CITATION

Cofini V, Piccardi L, Benvenuti E, Di
Pangrazio G, Cimino E, Mancinelli M,
Muselli M, Petrucci E, Picchi G, Palermo P,
Tobia L, Barbonetti A, Desideri G, Guido M,
Marinangeli F, Fabiani L and
Necozone S (2025) The I-KAPCAM-AI-Q: a
novel instrument for evaluating health care
providers' AI awareness in Italy.
Front. Public Health 13:1655659.
doi: 10.3389/fpubh.2025.1655659

COPYRIGHT

© 2025 Cofini, Piccardi, Benvenuti,
Di Pangrazio, Cimino, Mancinelli, Muselli,
Petrucci, Picchi, Palermo, Tobia, Barbonetti,
Desideri, Guido, Marinangeli, Fabiani and
Necozone. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

The I-KAPCAM-AI-Q: a novel instrument for evaluating health care providers' AI awareness in Italy

Vincenza Cofini¹, Laura Piccardi^{2,3}, Eugenio Benvenuti¹,
Ginevra Di Pangrazio¹, Eleonora Cimino¹, Martina Mancinelli⁴,
Mario Muselli^{1*}, Emiliano Petrucci⁵, Giovanna Picchi^{1,6},
Patrizia Palermo¹, Loreta Tobia¹, Arcangelo Barbonetti¹,
Giovambattista Desideri⁷, Maurizio Guido^{1,8},
Franco Marinangeli¹, Leila Fabiani¹ and Stefano Necozone¹

¹Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy,

²Department of Psychology, Sapienza University of Rome, Rome, Italy, ³IRCCS San Raffaele Cassino, Cassino, Italy, ⁴Department of Anatomical and Histological Sciences, Legal Medicine and Orthopedics, Sapienza University of Rome, Rome, Italy, ⁵Department of Anesthesia and Intensive Care Unit, San Salvatore Academic Hospital of L'Aquila, L'Aquila, Italy, ⁶Infectious Disease Department, Santa Rosa Hospital, Viterbo, Italy, ⁷Department of Medical and Cardiovascular Sciences, Sapienza University of Rome, Rome, Italy, ⁸Department of Pharmacy and Health Sciences and Nutrition, University of Calabria, Arcavacata di Rende, Italy

Background: Understanding healthcare providers' readiness and attitudes is crucial for integrating AI in healthcare, yet no validated tool exists to evaluate these aspects among Italian physicians. This study developed and validated the Italian Knowledge, Attitudes, Practice, and Clinical Agreement between Medical Doctors and the Artificial Intelligence Questionnaire (I-KAPCAM-AI-Q).

Methods: This was a cross-sectional validation study. The validation process included expert review ($n = 18$), face validity assessment ($n = 20$), technical implementation testing, and pilot testing ($n = 203$) with both residents and specialists. The questionnaire contained 29 items, one clinical universal scenario, and 6 clinical scenarios specific to 6 specialists.

Results: The questionnaire demonstrated strong content validity (S-CVI/Ave = 0.98) and acceptable internal consistency (Cronbach's Alpha = 0.7481, KR-21 = 0.832). Pilot testing revealed only 17% of participants had received digital technology training during medical education, while 91% showed clinical agreement with AI-proposed diagnoses. Knowledge in diagnostics was highest among AI applications (48%). Residents showed higher interest in technical support (58.3% vs. 42.0%, $p = 0.021$) and evidence-based validation (61.2% vs. 47.0%, $p = 0.043$) compared to specialists.

Conclusion: The I-KAPCAM-AI-Q provides a reliable tool for assessing healthcare providers' AI readiness and highlights the need for enhanced digital health education in medical curricula.

KEYWORDS

healthcare digitalization, questionnaire validation, medical AI literacy, medical technology integration, artificial intelligence awareness, medical technology acceptance, crossgenerational medical education

1 Introduction

Since the advent of artificial intelligence (AI) in 1955, its applications have rapidly expanded within healthcare, driven by both technological progress and growing societal expectations (1). In medical education, AI can enable innovative approaches by recording teaching videos, facilitating distance learning, managing resources, and supporting virtual inquiry systems (2). Additionally, AI can help process large repositories of medical data to enhance diagnostics and support problem-solving (3). It's important to distinguish between two levels of AI proficiency in healthcare: general AI recognition and clinical AI applications specific to a particular domain. In the digital age, a healthcare professional should have general AI awareness, encompassing fundamental understanding of AI concepts, capabilities, and limitations. This involves comprehending machine learning principles, data requirements, and ethical issues. In contrast, domain-specific clinical AI applications refer to the practical implementation of AI tools within specific medical contexts, such as AI-assisted diagnostic imaging in radiology, predictive models in intensive care, or natural language processing for clinical documentation (4).

AI in medicine applies computational techniques to enhance healthcare outcomes, clinical workflows, and medical research (5), offering solutions to complex healthcare challenges (6). These applications range from general-purpose tools adapted for healthcare (such as large language models like ChatGPT) to highly specialized clinical AI systems designed for specific medical tasks. Its evolving capabilities demonstrate exponential growth in diagnostic precision (7, 8), with domain-specific AI-driven image recognition achieving 25% higher accuracy than traditional methods in detecting early-stage anomalies (9, 10).

Yet harnessing this potential requires confronting human barriers. Recent studies revealed a significant lack of AI literacy among medical trainees: they often lack hands-on experience and a strong foundational knowledge of its application. Integrating AI into medical diagnostic workflows speeds up medical data analysis (11–13). The VALIDATE Project's findings are particularly striking for Italy, where over 80% of young physicians lack adequate training in digital health technologies crucial for healthcare modernization (14). While AI allows systematic storage and management of medical data through electronic health records (EHRs), cloud storage, and natural language processing (NLP) for data categorization (15), this technical infrastructure cannot compensate for deficits in human competencies. The lack of AI education represents a significant obstacle to effectively integrating AI-powered tools and technologies into clinical practice. Examining attitudes and perceptions of medical professionals, training physicians and AI proposals is essential for addressing this issue.

The challenges go beyond technical skills and include fundamental ethical considerations as well. Unlike human practitioners, AI systems lack moral agency and the capacity for value-based judgment (16, 17), raising important questions about their role in clinical decision-making. These concerns are amplified by the rapid adoption of tools like ChatGPT, originally developed as general-purpose language models but increasingly employed for medical information retrieval (18)—despite not being specifically designed for healthcare applications.

Italy's substantial PNRR investments in AI (19) contrast sharply with the VALIDATE findings on physician unpreparedness (14). AI

presents significant opportunities for integrated home care services, particularly relevant for older adults (20). The Italian National Health Service prioritizes home care to address aging demographics and reduce hospital pressure, resulting in cost savings (21). However, only 3% of Italians aged 65 + received home assistance in 2021, significantly lower than Northern European countries (22). The pandemic has further emphasized the importance of patient telemonitoring in both facilities and homes (23).

Italy's healthcare system already employs Clinical Decision Support Systems (CDSS) including Clinical Risk Prediction Models (CRPMs) for conditions like heart disease and sepsis (18), and Computerized Physician Order Entry (CPOE) systems for prescription management (24). These represent domain-specific clinical AI applications that require both general AI understanding and specialized clinical knowledge for effective use. Despite accelerating adoption and proven benefits in reducing medical errors, implementation remains variable across regions, highlighting the need for standardized physician competency assessment.

To systematically address these disparities, we developed the I-KAPCAM-AI-Q questionnaire, which represents the first validated Italian instrument designed to:

Evaluate healthcare professionals' knowledge and attitudes toward AI in clinical practice (5, 7, 8, 14).

Identify key barriers and facilitators of AI integration (16, 17, 19, 25).

Assess diagnostic agreement between physicians and AI systems using real-world scenarios (9, 10, 18, 24, 25).

Examine differences between medical residents and specialists (26–29).

2 Methods

2.1 Study design and setting

This cross-sectional validation study developed and validated the I-KAPCAM-AI-Q questionnaire, the first instrument of its kind specifically designed to evaluate the awareness, knowledge, attitudes, practices, and clinical agreement of Italian physicians regarding artificial intelligence. The study uses a mixed-methods research approach, which combines quantitative and qualitative elements, to validate the questionnaire and collect in-depth data. Specifically, the clinical scenarios are based on real-world situations, developed by clinicians, and reviewed by experts to ensure their relevance. These scenarios were then submitted to ChatGPT [ChatGPT (OpenAI, 2024). ChatGPT (Version GPT-3.5). [AI language model]. <https://openai.com>], and participants were asked to assess their agreement with the AI-generated diagnoses, which could include the correct diagnosis for comparison. The I-KAPCAM-AI-Q is important to assess the comprehension of AI's role in diagnostic processes and to measure trust in AI-generated recommendations. Moreover, it provides current usage patterns of AI tools in clinical practice, allows identifying barriers and facilitators to AI adoption and measures willingness to incorporate AI into daily clinical workflows. This last point is crucial in adhering to the current Government guidelines regarding digital healthcare.

Before questionnaire development, we conducted a comprehensive literature review was conducted, searching articles from multiple databases, including PubMed, Scopus, and Web of Science. This

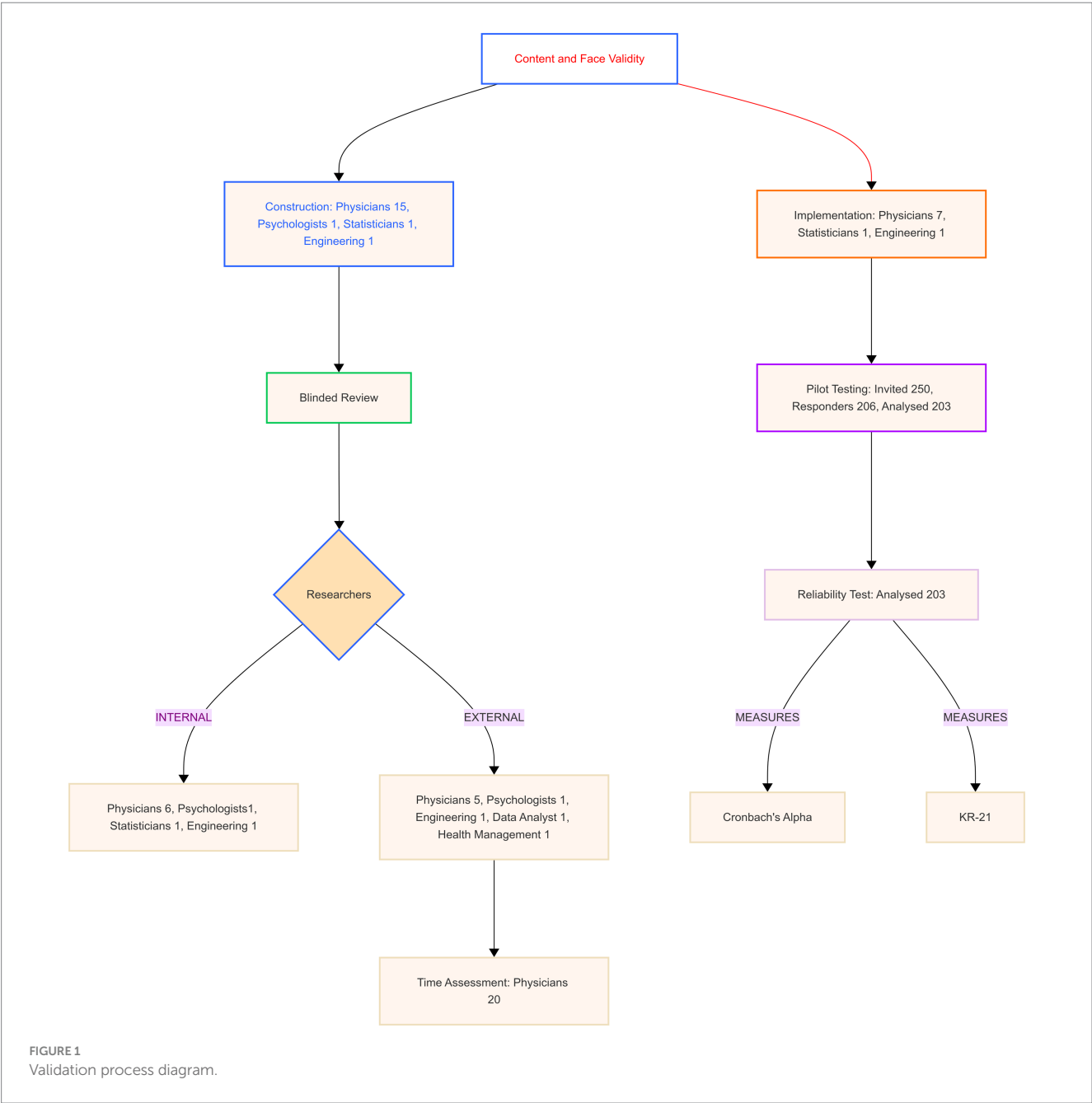
review focused on physician and medical student acceptance of clinical AI, helping identify key themes and knowledge gaps in the field and informing the subsequent questionnaire development. In particular, the review focused on survey or cross-sectional studies with attention to the validation test reported (12, 30–33). A crucial aspect, considering that the performance evaluation of large language models (LLMs) like ChatGPT in medical contexts is becoming an increasingly relevant area of research (34, 35), and such research, while focusing mainly on performance, highlights the importance of comprehensively analyzing the implications of using such tools, including the perspectives and concerns of healthcare professionals.

Figure 1 shows the steps in developing the Italian Knowledge, Attitudes, Practice, and Clinical Agreement between Medical Doctors and the Artificial Intelligence Questionnaire (I-KAPCAM-AI-Q).

2.2 Participants

The validation process involved multiple groups of participants, as reported in Figure 1. All participants gave their consent to participate, and the subgroups consisted of the following:

- Three medical doctors (two residents in hygiene and preventive medicine and one practicing physician), one statistician, and one engineer for the initial design;
- Ten specialized physicians representing diverse fields (hygiene and preventive medicine, endocrinology, infectious diseases, obstetrics and gynaecology, anesthesiology, occupational health, and geriatrics) for the development of clinical scenarios; The clinical domains were chosen to ensure representative coverage



of key medical specialties and thematic areas relevant to daily clinical practice, encompassing preventive medicine, acute care, chronic disease management, women's health, and occupational medicine for comprehensive assessment of AI integration potential;

- A panel of expert reviewers for content validation coming from different institutions and with roles ranging from clinical practice to research and teaching;
- Volunteer physicians not involved in the research for time assessment;
- Volunteer physicians not involved in the research for face validity assessment;
- A sample of physicians for pilot testing, with analysis focusing on differences between specialization groups.

2.3 Validation process

The validation methodology consisted of four sequential steps:

2.3.1 Step 1: construction and initial validation content validity

The questionnaire was organized into two parts to evaluate general perspectives and clinical reasoning comprehensively.

2.3.1.1 First part: general sections

The first part of the questionnaire consisted of six structured sections with items structured on answers yes/no or categorical or on a five-point Likert scale, and the last section for Open-ended Feedback:

- 1 Demographic and Work Information: this section collected data on respondents' medical specialty, years of experience, and previous familiarity with digital technologies (Number of items: 9).
- 2 Knowledge about AI Technology in Medicine: this section assessed respondents' knowledge of AI concepts and AI tools in clinical settings (Number of items: 3: 1-A question about participation in an AI course, designed to capture prior educational exposure; 2—A self-assessed level of knowledge question; – 3 A set of dichotomous questions to assess factual knowledge about specific AI concept plus 2 optional items linked to advanced knowledge declared).
- 3 AI Use and Experience in Medicine: this section focused on using AI technologies in clinical practice (Number of items: 3).
- 4 Attitudes and Awareness Regarding AI in Medicine: this section evaluated participants' viewpoints on AI (Number of items: 7 on a five-point Likert scale).
- 5 Willingness to Use AI in Medicine: this section explored participants' openness to integrating AI into their clinical workflows (Number of items: 3, with one item on a five-point Likert scale).
- 6 Willingness to Learn AI in Medicine: this section assessed participants' interest in gaining knowledge and skills related to AI (Number of items: 2).
- 7 Open-ended Feedback: in this section, participants were invited to provide recommendations, suggestions, or general feedback about the questionnaire or the topics covered.

The validation content validity of the first part of the I-KAPCAM-AI-Q, was assessed through expert evaluation based on four key attributes: relevance, clarity, simplicity, and ambiguity. We utilized The Content Validity Index (CVI) was used to quantify the evaluation process, with items scoring above 0.79 considered sufficiently relevant (32). Expert reviewers rated each item on a four-point scale from “not relevant” to “highly relevant.” Items scoring below the threshold were modified to improve clarity or alignment with study objectives. For face validity, volunteer physicians evaluated the questionnaire and were encouraged to seek clarification from the research team via telephone or email when needed. They documented concerns about specific questions while assessing response accuracy, comprehension of technical terminology, scenario realism, and completion time.

2.3.1.2 Second part: clinical scenarios

The second part of the questionnaire included two categories of clinical scenarios: “universally applicable scenarios” and “specialty-specific scenarios,” aimed at evaluating the level of concordance between the respondents' answers and ChatGPT's diagnostic proposals. These specialty-specific scenarios were presented only to participants who had declared the corresponding specialty.

The research team developed clinical scenarios and submitted them to AI for differential diagnoses and recommendations. All scenarios were based on real clinical cases with confirmed diagnoses and went through a structured validation process. Clinical experts from seven medical specialties reviewed the scenarios, considering clinical relevance (common presentations in clinical practice), appropriate complexity level, and ensuring clarity and completeness of the clinical information presented (36).

For each scenario, a structured presentation of the clinical case was provided to ChatGPT. The clinical scenarios were submitted to ChatGPT using the question: “What diagnosis would the patient receive based on the information provided?” (36).

The final tool included:

- 1 Universally applicable scenario:

Upon completing the first general sections, all participants were presented with a clinical scenario designed to be universally applicable to medical graduates. This scenario was simple, ensuring that even newly qualified physicians with limited clinical experience could engage meaningfully. Its simplicity enabled consistent participation across varying levels of expertise, while providing valuable insights into general medical reasoning and decision-making skills.

- 2 Speciality-specific scenarios:

The questionnaire also included seven distinct clinical scenarios tailored to the following specialties (Hygiene and Preventive Medicine, Infectious Diseases, Obstetrics and Gynecology, Anesthesiology, Geriatrics, Occupational Health, and Endocrinology).

The questionnaire was used to gather responses from participants to assess their level of agreement with AI's proposals, using a 5-point Likert scale ranging from ‘strongly disagree’ to ‘strongly agree’.

The structure of the questions and the response formats were identical across all scenarios. All participants were asked the same question, and the response structure provided by ChatGPT was consistent for each scenario.

The main purpose of these scenarios was to provide context for specific constructs and behaviors, not to measure a latent variable, so psychometric validation was not considered necessary.

The initial validation study centered on analyzing the responses to the universally applicable scenario, which was created to be answered by all participants. Despite being included in the final validated tool, the specialty-specific scenarios were not analyzed during this phase, because the pilot sample had a small number of specialists per category.

At the end of the questionnaire, participants could report how they had the link to the questionnaire.

2.3.2 Step 2: time assessment and questionnaire implementation

Twenty expert readers (medical doctors) were enrolled to evaluate the time to complete the questionnaire. The time collected was related to the total time spent to fill all questionnaire items and the time spent in each section and clinical scenarios designed. This activity was also useful for verifying and validating participant data.

The Questionnaire was designed for future web-based research studies, and technical aspects were evaluated, including:

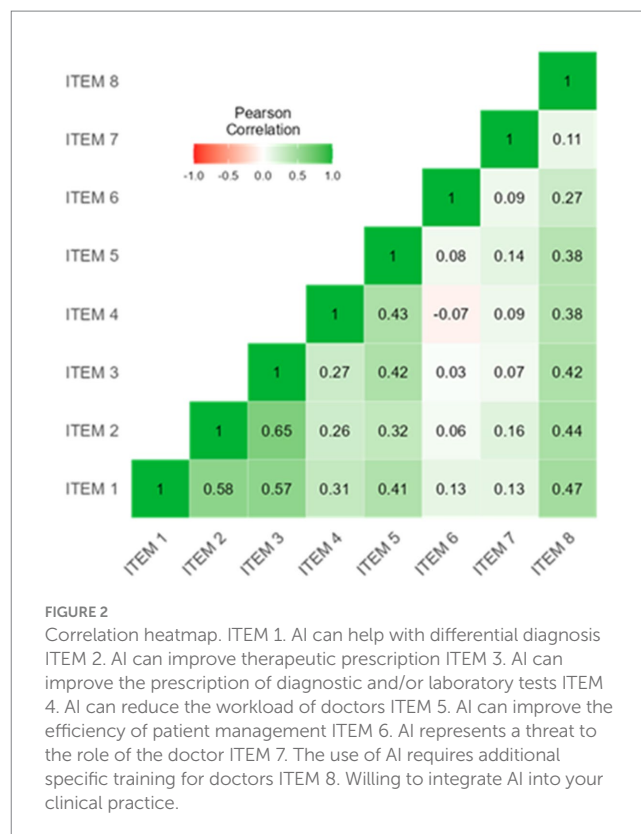
- o Survey platform capabilities, such as adaptive skip logic implementation and real-time progress saving functionalities.
- o Mobile device compatibility and cross-platform devices to ensure accessibility and usability.
- o Advanced data export functionalities for subsequent analysis.
- o User interface optimization, including dynamic progress indicators and clearly defined section transitions for enhanced respondent experience.
- o Integrated help text for the clarification of technical terminology, improving respondent comprehension.
- o Error prevention mechanisms to mitigate data entry inconsistencies and enhance response accuracy.
- o Comprehensive data management and security protocols, including encrypted data storage, GDPR-compliant handling (37), and secure export functionalities. Data management and security (data export functionality, encryption, GDPR-compliant data handling).
- o Strict adherence to accessibility standards, ensuring full compliance with Web Content Accessibility Guidelines 2.1 (WCAG 2.1) (38) for inclusive digital survey deployment. Accessibility standards (full adherence to WCAG 2.1 guidelines).

2.3.3 Step 3: pilot testing study

An online pilot test was conducted to investigate potential response differences between specific subgroups, particularly focusing on their medical specialization status. We analysed categorical variables using percentages and continuous variables using means and standard deviations. Group comparisons were performed using Chi-square, Fisher Exact, and Mann-Whitney U tests. Data were collected from July to September 2024.

2.3.4 Step 4: reliability assessment internal consistency

For this step, Cronbach's Alpha and Kuder-Richardson Coefficient (KR21) were used. Cronbach's Alpha for Likert-scale items, particularly the eight items measuring opinions and willingness to use



AI in medicine. Two items were reverse-coded for consistent directional interpretation. Values ≥ 0.90 were interpreted as excellent, 0.80–0.89 as good, and 0.70–0.79 as acceptable reliability (39). Item-level statistics included item-test correlation, item-rest correlation, and the impact of item removal on Cronbach's Alpha to evaluate individual item contributions to overall reliability. Kuder-Richardson Coefficient (KR21) was used for dichotomous data or all data split into categories. KR21 with values ≥ 0.80 indicates good to excellent reliability, and 0.70–0.79 is acceptable (40). Internal consistency analysis was not performed for each dimension separately, as each dimension included both Likert-scale and dichotomous questions. The instrument was developed based on a well-established theoretical framework, so we followed COSMIN guidelines by using content validity and internal consistency metrics (41). Therefore, we did not conduct exploratory or confirmatory factor analysis. Analysis was performed using STATA 18 BE, setting alpha 0.05. Figure 1 was made using the Mermaid Diagramming and charting tool,¹ and Figure 2 was from Jamovi 2.6.13.

3 Results

The results presented in this section describe the validation process of the I-KAPCAM-AI-Q questionnaire through four sequential steps, each designed to evaluate a specific aspect of the instrument. The data collected during pilot testing provided

¹ <https://mermaid.js.org/>

interesting preliminary insights into AI perceptions in medicine. They were primarily used to verify the effectiveness and reliability of the assessment tool and should not be interpreted as representative of the general population of Italian physicians. The analysis of results therefore focuses on the psychometric properties of the questionnaire and its ability to detect significant differences between groups, thus demonstrating its utility as a research instrument.

3.1 Step 1: the I-KAPCAM-AI-Q validation

As reported in [Figure 1](#), a panel of 18 experts from relevant medical or non-medical fields was convened to review the tool. For each item, I-CVI > 0.88 S-CVI/Ave (Average Approach = 0.98) and S-CVI/UA (Universal Agreement Approach = 0.82). The content validity ratio (CVR) ranges from 0.78 to 1.

3.2 Step 2: time assessment and implementation

To evaluate the practical feasibility and usability of the questionnaire, 20 physicians assessed the time required to complete the online questionnaire.

The average completion time for the questionnaire with the universal clinical scenario was 7.74 min was 7.74 minutes (SD = 3.65). The time spent completing the specialised clinical scenarios is reported in [Table 1](#).

3.2.1 Technical implementation results

After evaluating open-source survey platforms, LimeSurvey was selected as the optimal solution due to its comprehensive features and user-friendly interface ([42](#), [43](#)). The platform demonstrated capabilities aligned with our technical requirements, particularly in handling complex survey logic and ensuring data security compliance.

3.2.2 System implementation

The survey data collection was deployed on a dedicated hosting environment, integrating LimeSurvey with a high-performance LAMP stack (Linux ([44](#)), Apache ([45](#)), MySQL ([46](#)), PHP ([47](#))). The system architecture was designed to ensure high availability, data integrity, and enhanced security measures through:

- o Deployment of the latest stable LimeSurvey version.
- o MySQL was implemented as the database management system, supporting data storage and retrieval.
- o Implementation of file and directory permission settings to strengthen system security and prevent unauthorized access.
- o Configuration of survey parameters and interface language settings Implementation of user interface optimization, including dynamic progress indicators and clearly defined section transitions for enhanced respondent experience.
- o Implementation of industry-standard security protocols, including Transport Layer Security (TLS) encryption ([48](#)), Secure Sockets Layer (SSL) certification ([49](#)), and HyperText Transfer Protocol over Secure Socket Layer (HTTPS) ([48](#)) enforcement for secure data transmission.

3.2.3 Technical features and functionality

The implemented system successfully demonstrated several key functionalities:

- Survey Logic and Data Management:
- Robust skip logic handling with accurate participant pathway direction.
- Reliable progress-saving capability allowing session resumption.
- Multi-format data export, including Comma-separated values (CSV) ([50](#)), Excel Open XML Spreadsheet (XLSX) ([51](#)).
- Secure data transmission and storage compliant with data protection regulations.

3.2.4 Cross-platform compatibility

- Verified functionality across web browsers (Google Chrome, Mozilla Firefox, Apple Safari).
- Responsive design for mobile devices (iOS and Android).
- Consistent display of all survey elements regardless of device.

3.2.5 User interface elements

- Dynamic progress indicators accurately reflecting completion status.
- Clear section delineation with descriptive headers.
- Context-sensitive help text for technical terms.
- Error prevention mechanisms including response validation.
- Intuitive navigation between sections.

3.2.6 Pilot testing and expert review

The technical implementation was validated through two phases:

- 1 Initial functionality testing to verify LimeSurvey's operational status.
- 2 Pilot testing with nine experts from relevant medical and non-medical fields, as shown in [Figure 1](#), who assessed both usability and technical performance.

3.3 Step 3. pilot testing study

A pilot study was conducted on 206 responders (203 completed questionnaires analyzed), to assess the questionnaire's ability to capture meaningful variations between different groups of physicians and to identify potential improvements needed. [Table 2](#) reports on the characteristics of the participants. The average age was 41 (SD = 14), approximately 51% were residents, and 43% were identified as men.

The questionnaire's ability to differentiate knowledge levels and detect variations between professional groups was assessed through the analysis of responses to the AI knowledge section, as shown in [Table 3](#). The results demonstrate the instrument's sensitivity in capturing different levels of AI knowledge and experience.

The study evidenced that 17% of the sample did have specific training in digital or information technologies during their medical course, 22% attended courses or seminars on AI in

medicine, and 21% declared “no knowledge of AI.” Among the AI applications, the sample reported that Knowledge in diagnostics was higher (49%), and there were no significant differences between groups (residents versus specialists), as reported in Figure 3.

TABLE 1 Time assessment to fill the questionnaire by medical speciality.

Specialty	Physicians	Meantime (Minutes)	Standard deviation (Minutes)
Hygiene and Preventive Medicine	5	10.13	3.4
Infectious Diseases	2	11.91	10.3
Obstetrics and Gynecology	4	7.87	3.4
Anesthesiology	3	7.55	0.4
Geriatrics	5	10.76	7.8
Occupational Health	4	9.00	4.1
Endocrinology	5	6.19	1.1

TABLE 2 Participant characteristics (n = 203).

Characteristics	N (%) or mean (SD)
Sex	
Male	88 (43.3%)
Female	113 (55.7%)
Other	2 (1.0%)
Age	41.2 (14.1)
Specialization status	
Residents	103 (50.7%)
Specialists	100 (49.3%)
Years workers	12.4 (13.2)
Did you have any specific training in digital or information technologies during your studies?	
Yes	35 (17%)
No	168 (83%)

Overall, only 19.2% of participants reported using AI tools in clinical practice, with no significant differences between residents (17.5%) and specialists (21.0%) ($p = 0.524$). The most used applications were diagnostic imaging (6.9%), chatbots for diagnostic queries (8.4%), and clinical decision support systems (5.4%). Other applications, such as machine learning in diagnosis (2.5%) and electronic health record analytics (3.4%), had very low adoption rates. At the same time, tools like drug discovery, genome editing, and robotic surgery were nearly unused. No significant differences were observed between the groups for all applications (all $p > 0.05$), indicating a generally low level of AI integration into clinical practice (Figure 4).

Table 4 displays the results of the instrument's effectiveness evaluation in measuring attitudes and awareness using a five-point Likert scale. Participants demonstrated consistently positive attitudes toward AI's role in medicine, with high mean scores across groups for its potential in differential diagnosis (mean 4.9, SD 0.6) and improving therapeutic prescriptions (mean 4.6, SD 0.8). No significant differences were observed between residents and specialized participants in any category ($p > 0.05$). The strong consensus was on the need for additional training in AI for doctors (mean 5.4, SD 0.5), highlighting a key area for curriculum development. Concerns about AI posing a threat to the doctor's role were moderate (mean 3.4, SD 0.9).

The analysis presented in Figure 5 evaluated the questionnaire's capacity to assess willingness to adopt AI and identify barriers and facilitators. Most respondents expressed a positive attitude toward integrating AI into clinical practice, with 26.6% indicating they were “very available” and 53.2% stating they were “available.” There were no significant differences observed between residents and specialists ($p = 0.757$). The primary barriers identified included a lack of training (80.3%) and resistance to change (53.7%), with no significant differences between the two groups. Additionally, 38.4% of participants voiced concerns about privacy and data security. Residents were more likely to report a lack of scientific evidence regarding the effectiveness (34.0%) compared to specialists (21.0%) ($p = 0.039$). Education and professional development (87.7%) were the most endorsed incentives for integrating AI and ongoing technical support (50.2%). Notably, residents emphasized the need for continuous technical support more than specialists, with 58.3%

TABLE 3 Knowledge about AI Technology in medicine.

Items	Specialization status			p-value
	Residents (N = 103)	Specialists (N = 100)	Total	
	N (%)	N (%)	203 (100.0%)	
How would you describe your general knowledge of AI?				
None	19 (18.4%)	25 (25.0%)	44 (21.7%)	0.310
Basic knowledge	70 (68.0%)	60 (60.0%)	130 (64.0%)	
Intermediate knowledge	14 (13.6%)	13 (13.0%)	27 (13.3%)	
Advanced knowledge	0 (0.0%)	2 (2.0%)	2 (1.0%)	
Have you attended any courses or seminars on AI in medicine?				
No	84 (81.6%)	82 (82.0%)	166 (81.8%)	0.934
Yes	19 (18.4%)	18 (18.0%)	37 (18.2%)	

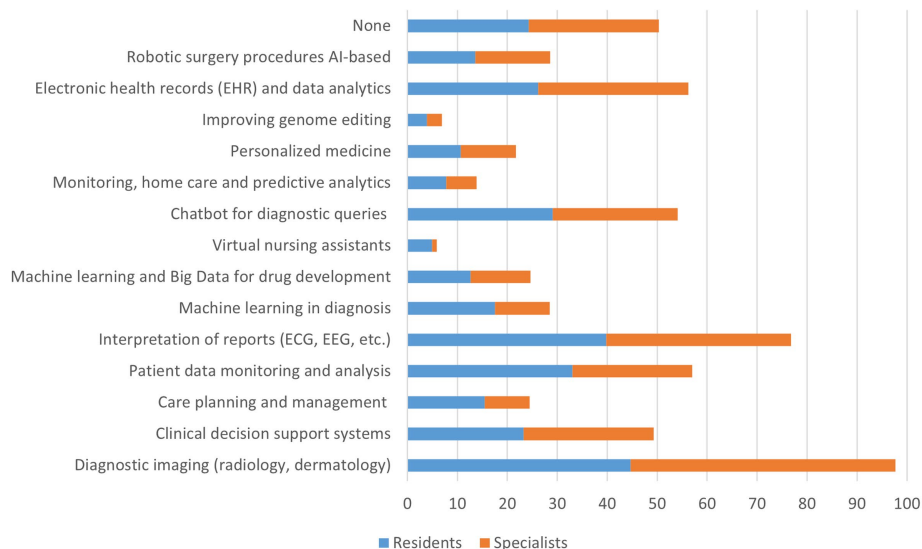


FIGURE 3
Answers to the question “Which of the following AI applications do you know?”.

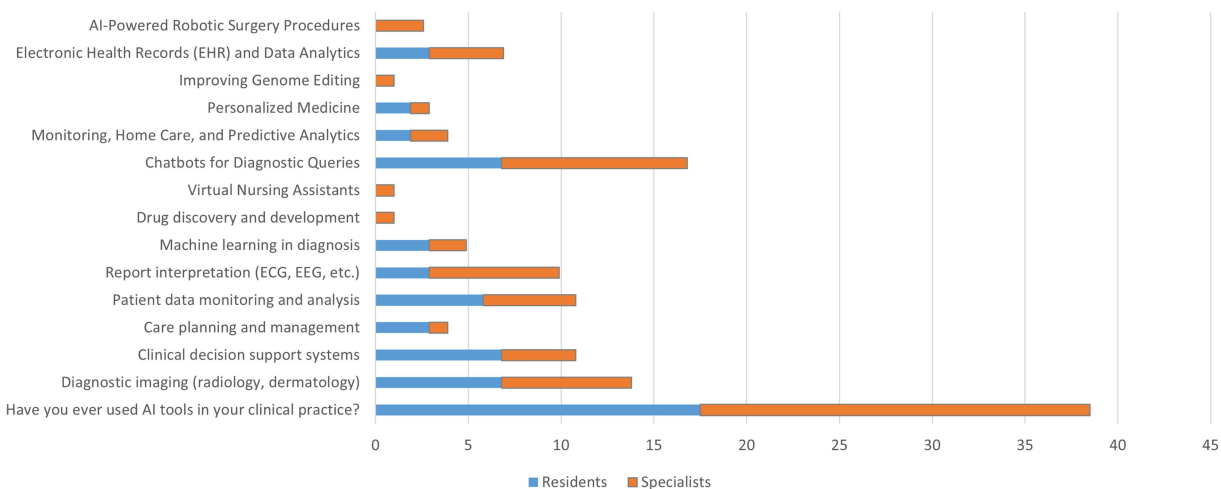


FIGURE 4
AI Use and experience in medicine in the last year.

of residents versus 42.0% of specialists expressing this view ($p = 0.021$). Furthermore, 54.2% of participants highlighted the importance of scientific evidence regarding efficacy and safety, with a higher percentage of residents (61.2%) compared to specialists (47.0%) recognizing this need ($p = 0.043$).

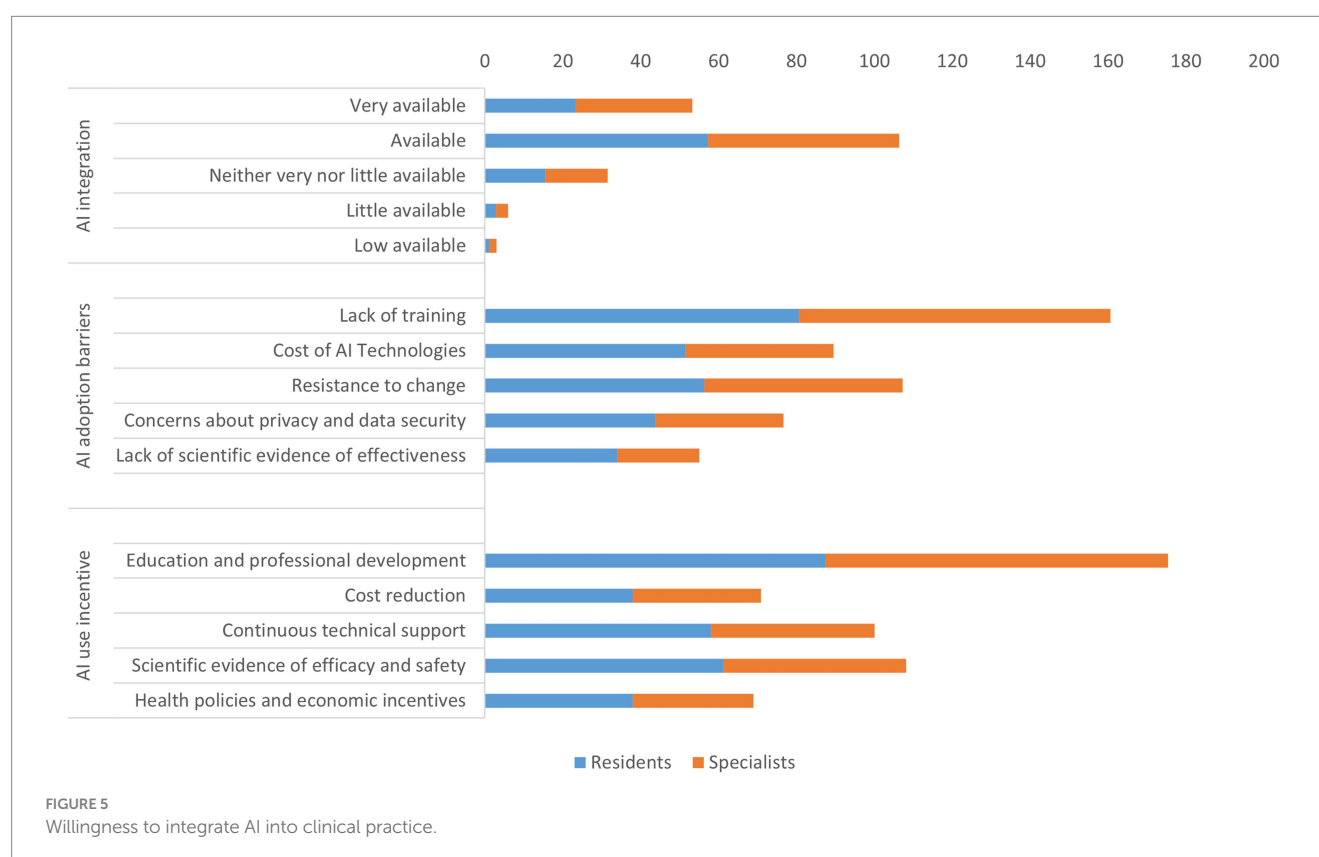
Assessing willingness to engage in artificial intelligence medical training, the analysis demonstrated overwhelming interest in AI education, with 95.1% of respondents expressing willingness to participate in training programs (Table 5). Both career stages showed a consistent level of interest, with 93.2% of residents and 97.0% of specialists indicating interest, and there were no significant differences ($p = 0.211$).

Regarding preferred training modalities, online courses emerged as the most popular format, with 61.1% of all respondents favouring this option. Residents showed a somewhat higher preference for online learning (67.0%) compared to specialists (55.0%), though this difference did not reach statistical significance ($p = 0.080$). The second most preferred modality was hands-on training, with 60.6% of respondents choosing it, with almost identical preferences among residents (60.2%) and specialists (61%), $p = 0.906$.

In-person seminars and workshops were preferred by 45.3% of respondents, with specialists showing greater interest (51.0%) than residents (39.8%; $p = 0.109$). Webinars and conferences were the least

TABLE 4 Attitudes and awareness regarding AI in medicine.

Items	Residents (N = 102)	Specialised (N = 99)	Total (N = 201)	p-value
	Mean (SD)	Mean (SD)	Mean (SD)	
AI can help with differential diagnosis	4.9 (0.6)	4.9 (0.7)	4.9 (0.6)	0.837
AI can improve therapeutic prescription	4.5 (0.9)	4.7 (0.7)	4.6 (0.8)	0.393
AI can improve the prescription of diagnostic and/or laboratory tests	4.7 (0.77)	4.7 (0.7)	4.7 (0.7)	0.665
AI can reduce the workload of doctors	4.7(0.8)	4.6 (1.0)	4.6 (0.9)	0.485
AI can improve the efficiency of patient management	4.6 (0.7)	4.6 (0.9)	4.6 (0.8)	0.795
AI represents a threat to the role of the doctor	3.5 (0.9)	3.4 (0.9)	3.4 (0.9)	0.646
The use of AI requires additional specific training for doctors	5.4 (0.5)	5.5 (0.5)	5.4 (0.5)	0.534



preferred format, with only 29.6% of respondents expressing interest in this modality. This preference was consistent across both groups, with 29.1% of residents and 30.0% of specialists selecting this option ($p = 0.892$).

The absence of statistically significant differences in preferences between residents and specialists suggests that career stage may not be determining the choice of AI learning formats.

3.4 Universally applicable scenario

The instrument's ability to measure clinical decision-making alignment with AI was assessed by validating clinical scenarios.

The universal scenario was structured as follows:

"A 32-year-old Caucasian patient has reported asthenia and evening fever for about 10 days. In medical history: cocaine addiction, history of stage IV Hodgkin lymphoma treated with chemotherapy in complete remission at the last follow-up in 2023, unprotected sexual intercourse, previous HPV infection. During the physical exam, it was determined that there were inguinal lymphadenopathy and a single genital ulcerative lesion in the balano-preputial area."

For this scenario, participants were asked to rate their feelings towards three diagnoses: the first AI-generated diagnosis and two other AI proposals generated for other clinical scenarios.

In the general scenario, clinical agreement with the diagnosis among those proposed by AI was 91% (among specialist doctors: 92%), as reported in Table 6. The consistent response patterns across

TABLE 5 Willingness to learn AI in Medicine.

Items	Specialization status			<i>p</i> -value
	Residents (<i>N</i> = 103)	Specialists (<i>N</i> = 100)	Total (<i>N</i> = 203)	
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	
Would you be interested in participating in training programs on AI applied to medicine?				
No	7 (6.8%)	3 (3.0%)	10 (4.9%)	0.211
Yes	96 (93.2%)	97 (97.0%)	193 (95.1%)	
Which modalities would you prefer?				
Online courses				
No	34 (33.0%)	45 (45.0%)	79 (38.9%)	0.080
Yes	69 (67.0%)	55 (55.0%)	124 (61.1%)	
In-person seminars and workshops				
No	62 (60.2%)	49 (49.0%)	111 (54.7%)	0.109
Yes	41 (39.8%)	51 (51.0%)	92 (45.3%)	
Hands-on training				
No	41 (39.8%)	39 (39.0%)	80 (39.4%)	0.906
Yes	62 (60.2%)	61 (61.0%)	123 (60.6%)	
Webinars and conferences				
No	73 (70.9%)	70 (70.0%)	143 (70.4%)	0.892
Yes	30 (29.1%)	30 (30.0%)	60 (29.6%)	

TABLE 6 Agreement between physicians and AI in differential diagnosis.

Items	Specialization status			<i>p</i> -value
	Residents (<i>N</i> = 103)	Specialists (<i>N</i> = 100)	Total (<i>N</i> = 203)	
	<i>N</i> (%)	<i>N</i> (%)	<i>N</i> (%)	
Syphilis				
Strongly disagree	2 (1.9%)	2 (2.0%)	4 (2.0%)	0.902
Disagree	1 (1.0%)	0 (0.0%)	1 (0.5%)	
Neutral	7 (6.8%)	6 (6.0%)	13 (6.4%)	
Agree	68 (66.0%)	68 (68.0%)	136 (67.0%)	
Strongly agree	25 (24.3%)	24 (24.0%)	49 (24.1%)	
Irritable Bowel Syndrome (IBS)				
Strongly disagree	37 (35.9%)	31 (31.0%)	68 (33.5%)	0.823
Disagree	52 (50.5%)	54 (54.0%)	106 (52.2%)	
Neutral	11 (10.7%)	11 (11.0%)	22 (10.8%)	
Agree	3 (2.9%)	3 (3.0%)	6 (3.0%)	
Strongly agree	0 (0.0%)	1 (1.0%)	1 (0.5%)	
Pathology inflammatory pelvic				
Strongly disagree	15 (14.6%)	13 (13.0%)	28 (13.8%)	0.798
Disagree	36 (35.0%)	36 (36.0%)	72 (35.5%)	
Neutral	29 (28.2%)	24 (24.0%)	53 (26.1%)	
Agree	22 (21.4%)	24 (24.0%)	46 (22.7%)	
Strongly agree	1 (1.0%)	3 (3.0%)	4 (2.0%)	

different professional groups further support the reliability of this section of the questionnaire.

3.5 Step 4: internal consistency

The reliability analysis of the 8-item scale (opinion and willingness to use AI), yielded a Cronbach's Alpha of $\alpha = 0.7481$, indicating acceptable internal consistency as reported in Table 7. Item-level statistics showed that most items had moderate to high item-test correlations (Item 1 = 0.7474; Item 8 = 0.7178). However, two items, Item 6 and Item 7, had lower item-test correlations (0.3277 and 0.3713, respectively), suggesting weaker alignment with the overall construct.

The average interim correlation was 0.2707, within the acceptable range of measures assessing a unidimensional construct.

The interim correlation matrix (Figure 2) revealed moderate to strong correlations among most items (Item 1 and Item 2 = 0.58). In contrast, certain items, such as Item 6, exhibited weaker correlations with other items (Item 6 and Item 3 = 0.03).

Excluding specific items such as Item 6 and Item 7 led to an increase in Cronbach's Alpha (0.7779 and 0.7700, respectively), suggesting that these items may contribute less to the scale's overall reliability.

The internal consistency of the dichotomous items was assessed using Kuder–Richardson Formula 21 (KR-21), which is calculated as follows:

$$KR-21 = k / (k-1) \{1 - [M \cdot (k-M)] / (k \cdot \text{var})\}$$

Where:

- $k = 18$ (total number of items),
- $M = 10$ (mean of total scores),
- $\text{var} = 20.7044$ (variance of total scores).

Using this formula, the KR-21 coefficient was calculated to be $\rho = 0.832$.

4 Discussion

The validation of the I-KAPCAM-AI-Q marks a crucial advancement in comprehending and enhancing the integration of AI within Italian healthcare system. This aligns with the recommendations of the Italian National Health Care System for utilizing digital tools in delivering healthcare. As a validated instrument, the I-KAPCAM-AI-Q offers a valuable resource for informing policy decisions, guiding institutional training initiatives, and supporting strategic planning related to AI implementation in clinical practice. Moreover, it is useful in helping to track changes in attitudes and practices over time. Our findings reveal several key insights that warrant detailed discussion. The I-KAPCAM-AI-Q demonstrated robust psychometric properties, with strong content validity (S-CVI/Ave = 0.98) and acceptable internal consistency (Cronbach's Alpha = 0.7481). The KR-21 coefficient of 0.832 for dichotomous

items further supports the instrument's reliability. These metrics align with or exceed those of similar validated healthcare questionnaires, suggesting that the I-KAPCAM-AI-Q is a reliable tool for assessing AI readiness in healthcare settings. For such a reason, I-KAPCAM-AI-Q holds potential for use in experimental studies related to AI, such as longitudinal studies about the adoption of AI in healthcare. It allows for comparative analysis across different medical specialties and supports evidence-based approaches to AI implementation. While the data collected in this pilot study cannot be generalized due to the study's specific objective, they nonetheless highlight interesting findings that can be further developed through targeted research. The I-KAPCAM-AI-Q builds on previous research assessing AI literacy and attitudes in healthcare but introduces significant methodological advancements. A striking finding from our preliminary study is the substantial growth in digital technology training, with only 17% of participants reporting specific training during their medical education. This finding aligns with recent evidence from the VALIDATE Project, which highlighted insufficient preparation in digital health among Italian physicians. The deficit in formal training may partially explain why only 19% of participants reported using AI tools in their practice, despite demonstrating a high degree of interest and openness to integration.

The growing capabilities of AI tools further underscore the urgency of addressing this educational gap. Recent studies by Rodrigues Alessi et al. (34, 35) have shown that ChatGPT-3.5 and ChatGPT-4.0 outperform medical students on national assessments, with accuracy increasing from 68.4 to 87.2% between versions. These findings highlight the accelerating capabilities of AI in clinical reasoning, making the lack of structured training particularly concerning. Similar trends have been observed across Europe, where Mousavi Baigi et al. (28) found that only 15–22% of healthcare students received formal AI training. These results support calls for integrating structured AI competencies into undergraduate and postgraduate medical curricula, as emphasized in the literature (4).

The comparison between residents and specialists revealed interesting patterns. While both groups showed similar levels of AI knowledge and general attitudes, residents demonstrated significantly higher interest in technical support (58.3% vs. 42.0%, $p = 0.021$) and evidence-based validation (61.2% vs. 47.0%, $p = 0.043$). This difference suggests that newer medical professionals may be more attuned to the importance of systematic implementation and validation of AI tools, possibly reflecting evolving perspectives in medical education. AlZaabi et al. (29), reported similar generational differences in AI readiness among physicians, with younger practitioners showing greater technological adaptability. The use of AI in medical practice reveals distinct patterns between junior doctors and specialists, highlighting crucial considerations for medical education and implementation (52–55). Junior doctors typically demonstrate greater capacity to embrace new technologies, but possess less clinical experience to contextualise AI outputs, potentially leading to over-reliance on AI recommendations without the benefit of extensive clinical intuition. Specialists, drawing from years of practice, have developed robust mental models and decision-making frameworks that enable them to evaluate AI applications critically. While their established workflows may create initial resistance to AI adoption, their deep specialty knowledge allows them to precisely identify where

TABLE 7 Reliability statistics.

Items	Average					
	N	Sign correlation	Item test correlation	Item-rest	interitem	Alpha
Opinion and willing to use AI						
ITEM 1. AI can help with differential diagnosis	203	+	0.7474	0.6305	0.2373	0.6853
ITEM 2. AI can improve therapeutic prescription	203	+	0.7219	0.5963	0.2431	0.6922
ITEM 3. AI can improve the prescription of diagnostic and/or laboratory tests	203	+	0.7126	0.5840	0.2453	0.6947
ITEM 4. AI can reduce the workload of doctors	203	+	0.5548	0.3850	0.2814	0.7327
ITEM 5. AI can improve the efficiency of patient management	203	+	0.6591	0.5145	0.2575	0.7083
ITEM 6. AI represents a threat to the role of the doctor	203	–	0.3277	0.1259	0.3335	0.7779
ITEM 7. The use of AI requires additional specific training for doctors	203	+	0.3713	0.1734	0.3235	0.7700
ITEM 8. Willing to integrate AI into your clinical practice	203	–	0.7178	0.5909	0.2441	0.6933
Test scale					0.2707	0.7481

AI adds value versus where it might complicate existing processes (56, 57). Generational differences in AI adoption observed in our study highlight the evolving perspective of younger physicians.

The results indicate a preference for flexible learning options (online courses) and practical experience (hands-on training) over traditional educational formats. The rapid expansion of medical knowledge makes it difficult for individual practitioners to stay fully informed about all advancements (36). Our findings The analysis demonstrates concerning levels of trust in ChatGPT for medical diagnostics, with 8.4% of participants regularly using the chatbot for diagnostic inquiries. The proportion of participants using ChatGPT for diagnostic inquiries (8.4%) justifies our methodological decision to investigate the agreement between ChatGPT and participant responses regarding real-life medical scenarios. Despite its limitations, this finding has real-world clinical relevance, as healthcare consumers are already turning to ChatGPT for medical decision-making (58). The high clinical agreement (91%) with AI-proposed diagnoses in the universal scenario is noteworthy, suggesting that healthcare providers can effectively evaluate AI-generated clinical recommendations, despite limited formal training. This finding has important implications for future integration of AI in clinical decision support systems and highlights the potential for AI to complement rather than replace clinical judgment.

The validation process of the I-KAPCAM-AI-Q has three main methodological limitations: the lack of a test–retest reliability assessment and the potential selection bias stemming from our volunteer-based sampling approach. Although these limitations are specific to the validation methodology, the pilot study conducted as part of this process has yielded valuable insights that will inform future applications of the tool. Moreover, the study did not include an in-depth analysis of the regulatory and ethical frameworks governing the use of AI in clinical practice, an important area for future research (35, 59). Third, we did not perform exploratory or confirmatory factor analysis, as our primary aim was to establish content validity and reliability for this novel tool. While this limits insights into potential latent structures, the predefined domains were rigorously validated through expert review and pilot testing. Future research should employ factor analysis to verify dimensional structure in larger, diverse

samples. Future applications of the validated I-KAPCAM-AI-Q could reveal multiple important research priorities. While our pilot study was instrumental in refining and validating the questionnaire, its findings should be viewed primarily as supportive of the tool's development rather than as generalizable results. Future large-scale applications of the I-KAPCAM-AI-Q should focus on conducting longitudinal studies to evaluate how knowledge and attitudes of physicians towards AI evolve over time with increased exposure to AI technologies. The questionnaire could be implemented with more diverse clinical scenarios across different medical specialties, to better understand the tool's effectiveness in various healthcare contexts. Future research through the I-KAPCAM-AI-Q can investigate variations in responses across different healthcare settings, geographical regions, and levels of technological infrastructure and the relationship between measured AI readiness and actual implementation success in clinical settings. While I-KAPCAM-AI-Q was specifically developed and validated within the Italian healthcare system to reflect the educational, clinical, and technological context in which Italian physicians operate, we acknowledge that the structure and content of the instrument may hold relevance in other settings where artificial intelligence is becoming increasingly integrated into healthcare practice. Any future adaptation for use in different countries would require a rigorous cross-cultural validation process, following established guidelines this approach ensures that translated versions retain semantic, experiential, and conceptual equivalence while remaining culturally appropriate and context-sensitive. Furthermore, future research should explore how the I-KAPCAM-AI-Q can be adapted to assess AI readiness in different healthcare systems and cultural contexts while maintaining its psychometric properties. The tool's comprehensive assessment of technical competency and clinical judgment makes it particularly suitable for developing targeted educational interventions across Italian healthcare, from university-level training programs to workplace implementation strategies. Several distinctive strengths enhance the significance of the I-KAPCAM-AI-Q in the evolving landscape of AI assessment tools. A primary strength lies in its demonstrated effectiveness across different levels of medical expertise, with particular value for early-career

physicians and residents. The evaluation of AI readiness among young doctors is especially crucial as they represent the future of healthcare and are often at the forefront of technological adoption. Notably, the I-KAPCAM-AI-Q stands as the first and only Italian validated instrument that systematically compares clinical decision-making with ChatGPT responses using real-world scenarios. This unique feature provides unprecedented insights into the alignment between human and artificial intelligence, reasoning in authentic clinical situations. The use of real clinical cases rather than theoretical scenarios enhances the tool's practical relevance and validity. The tool's power lies in its ability to assess clinical agreement, not only between physicians and AI, but also among healthcare providers themselves. This capability opens valuable avenues for understanding how different clinicians approach similar cases and how their decision-making patterns align or diverge when using AI tools. These perspectives are essential for developing standardized approaches to integrating AI in clinical practice and identifying areas where additional training or support may be required. These findings, combined with the tool's robust psychometric properties, ensure that I-KAPCAM-AI-Q is a comprehensive and innovative framework for advancing AI education and implementation in medicine. The ability to simultaneously assess technical competency, clinical judgment, and agreement patterns makes it an essential instrument for guiding the future of AI integration in healthcare settings.

Future research should prioritize the development of longitudinal studies to monitor changes in AI readiness and usage over time, particularly in response to targeted educational interventions. A possible roadmap could include: (1) integrating the I-KAPCAM-AI-Q into AI-focused modules as a pre- and post-training assessment tool; (2) validating its sensitivity to change over time (longitudinal psychometric testing); and (3) evaluating the impact of AI education on actual clinical practice, using mixed-method approaches. Such studies would help define the most effective educational formats and inform curriculum design tailored to different stages of medical training.

In conclusion, the I-KAPCAM-AI-Q represents a rigorously validated tool to assess AI readiness among Italian healthcare providers, demonstrating robust psychometric properties and high clinical concordance with AI-generated diagnoses. Our findings highlight critical gaps in AI education, with only 17% of participants reporting formal training in digital technologies, consistent with broader European trends (14, 28). We recommend integrating the I-KAPCAM-AI-Q into medical curricula and continuous professional development programs, particularly within the framework of Italy's National Recovery and Resilience Plan (PNRR) for digital health transformation. Future applications should focus on longitudinal implementation studies and specialty-specific adaptations to monitor the evolving relationship between physicians and AI technologies. As the first Italian instrument enabling direct comparison between physician judgment and AI recommendations, this tool provides a valuable benchmark for both clinical practice and health policy development in the era of digital medicine.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

VC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. LP: Investigation, Validation, Writing – original draft, Writing – review & editing. EB: Data curation, Investigation, Software, Validation, Writing – original draft, Writing – review & editing. GP: Investigation, Validation, Writing – review & editing. EC: Investigation, Validation, Writing – original draft, Writing – review & editing. MMA: Investigation, Validation, Writing – original draft, Writing – review & editing. EPM: Investigation, Validation, Writing – review & editing. GP: Investigation, Validation, Writing – original draft. PP: Investigation, Writing – original draft. LT: Investigation, Validation, Writing – original draft. AB: Investigation, Validation, Writing – original draft. MG: Investigation, Validation, Writing – original draft. GD: Investigation, Validation, Writing – original draft. FM: Investigation, Validation, Writing – original draft. LF: Investigation, Validation, Writing – original draft. SN: Investigation, Methodology, Validation, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

Authors thank all participants to all activity in the current study.

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 Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Zhao H., Li G., Feng W. (2018). Research on application of artificial intelligence in medical education. In *Proceedings of the International Conference on Engineering Simulation and Intelligent Control*, Hunan, China, 9, 340–342.
- Mir MM, Mir GM, Raina NT, Mir SM, Mir SM, Miskeen E, et al. Application of artificial intelligence in medical education: current scenario and future perspectives. *J Adv Med Educ Prof.* (2023) 11:133–40. doi: 10.30476/JAMP.2023.98655.1803
- Han ER, Yeo S, Kim MJ, Lee YH, Park KH, Roh H. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review. *BMC Med Educ.* (2019) 19:1–15. doi: 10.1186/s12909-019-1891-5
- Rabbani SA, El-Tanani M, Sharma S, Rabbani SS, El-Tanani Y, Kumar R, et al. Generative artificial intelligence in healthcare: applications, implementation challenges, and future directions. *BioMedInformatics.* (2025) 5:37. doi: 10.3390/biomedinformatics5030037
- Litjens G, Kooi T, Beijndorf BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005
- Tai MC. The impact of artificial intelligence on human society and bioethics. *Tzu Chi Med J.* (2020) 32:339–43. doi: 10.4103/tcmj.tcmj_71_20
- Cacciamani GE, Chu TN, Sanford DI, Abreu A, Duddalwar V, Oberai A, et al. PRISMA AI reporting guidelines for systematic reviews and meta-analyses on AI in healthcare. *Nat Med.* (2023) 29:14–5. doi: 10.1038/s41591-022-02139-w
- Firouzi F, Jiang S, Chakraborty K, Farahani B, Daneshmand M, Song J, et al. Fusion of IoT, AI, edge–fog–cloud, and Blockchain: challenges, solutions, and a case study in healthcare and medicine. *IEEE Internet Things J.* (2023) 10:3686–705. doi: 10.1109/JIOT.2022.3191881
- An Q, Rahman S, Zhou J, Kang JJ. A Comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges. *Sensors (Basel).* (2023) 23:4178. doi: 10.3390/s23094178
- Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. *Adv Exp Med Biol.* (2020) 123:–21. doi: 10.1007/978-3-030-33128-3_1
- Taylor R, Lawry GFT. Hacking healthcare: how AI and the intelligence revolution will reboot an ailing system. *Croat Med J.* (2022) 63:402–3. doi: 10.3325/cmj.2022.63.402
- Alowais SA, Alghamdi SS, Alsuehany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ.* (2023) 23:689. doi: 10.1186/s12909-023-04698-z
- Aamir A, Iqbal A, Jawed F, Ashfaq F, Hafsa H, Anas Z, et al. Exploring the current and prospective role of artificial intelligence in disease diagnosis. *Ann Med Surg.* (2024) 86:943–9. doi: 10.1097/MS9.0000000000001700
- Casà C, Marotta C, Di Pumpo M, Cozzolino A, D'Aviero A, Friscale EM, et al. COVID-19 and digital competencies among young physicians: are we (really) ready for the new era? *Ann Ital Super Sanita.* (2021) 57:1–6. doi: 10.4415/ANN_21_01_01
- Zhou SK, Greenspan H, Shen D. Deep learning for medical image analysis Elsevier (2023).
- Bottomley D, Thaldar D. Liability for harm caused by AI in healthcare: an overview of the core legal concepts. *Front Pharmacol.* (2023) 14:1297353. doi: 10.3389/fphar.2023.1297353
- Jason C, Amanda Z. Hey Watson – can i sue you for malpractice? Examining the liability of artificial intelligence in medicine. *Asia Pac J Health Law Ethics.* (2018) 11:51–80. doi: 10.38046/APJHLE.2018.11.2.004
- Shahmoradi L, Safdari R, Ahmadi H, Zahmatkeshan M. Clinical decision support systems-based interventions to improve medication outcomes: a systematic literature review on features and effects. *Med J Islam Repub Iran.* (2021) 35:27. doi: 10.47176/mjiri.35.27
- Agenzia nazionale per i servizi sanitari regionali. (2024) Mission 6 Health. Available online at: <https://www.agenas.gov.it/pnrr/missione-6-salute> (Accessed June 27, 2025).
- Cingolani M, Scendoni R, Fedeli P, Cembrani F. Artificial intelligence and digital medicine for integrated home care services in Italy: opportunities and limits. *Front Public Health.* (2023) 10:1095001. doi: 10.3389/fpubh.2022.1095001
- Dinesen B, Haesum LKE, Soerensen N, Nielsen C, Grann O, Hejlesen O, et al. Using preventive home monitoring to reduce hospital admission rates and reduce costs: a case study of telehealth among chronic obstructive pulmonary disease patients. *J Telemed Telecare.* (2012) 18:221–5. doi: 10.1258/jtt.2012.110704
- Istituto nazionale di statistica. (2021). Indagine sui presidi residenziali socio-assistenziali e socio-sanitari. Available online at: <https://www.istat.it/it/files/2021/03/12.pdf> (Accessed June 27, 2025).
- Shabbir A, Shabbir M, Javed AR, Rizwan M, Iwendi C, Chakraborty C. Exploratory data analysis, classification, comparative analysis, case severity detection, and internet of things in COVID-19 telemonitoring for smart hospitals. *J Exp Theor Artif Intell.* (2022) 35:507–34. doi: 10.1080/0952813X.2021.1960634
- Moja L, Liberati EG, Galuppo L, Gorli M, Maraldi M, Nanni O, et al. Barriers and facilitators to the uptake of computerized clinical decision support systems in specialty hospitals: protocol for a qualitative cross-sectional study. *Implement Sci.* (2014) 9:105. doi: 10.1186/s13012-014-0105-0
- Ferdush J, Begum M, Hossain ST. ChatGPT and clinical decision support: scope, application, and limitations. *Ann Biomed Eng.* (2024) 52:1119–24. doi: 10.1007/s10439-023-03329-4
- Raina A., Mishra P., Goyal H., Kumar D. (2024). AI as a medical ally: evaluating ChatGPT's usage and impact in Indian healthcare. arXiv:240115605.
- Chen M, Zhang B, Cai Z, Seery S, Gonzalez MJ, Ali NM, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. *Front Med.* (2022) 9:990604. doi: 10.3389/fmed.2022.990604
- Mousavi Baigi SF, Sarbaz M, Ghaddarpouri K, Ghaddarpouri M, Mousavi AS, Kimiafar K. Attitudes, knowledge, and skills towards artificial intelligence among healthcare students: a systematic review. *Health Sci Rep.* (2023) 6:e1138. doi: 10.1002/hsr.2.1138
- AlZaabi A, AlMaskari S, AalAbdulsalam A. Are physicians and medical students ready for artificial intelligence applications in healthcare? *Digit Health.* (2023) 9:20552076231152167. doi: 10.1177/20552076231152167
- Tamori H, Yamashina H, Mukai M, Morii Y, Suzuki T, Ogasawara K. Acceptance of the use of artificial intelligence in medicine among Japan's doctors and the public: a questionnaire survey. *JMIR Hum Factors.* (2022) 9:24680. doi: 10.2196/24680
- Pedro AR, Dias MB, Laranjo L, Cunha AS, Cordeiro JV. Artificial intelligence in medicine: a comprehensive survey of medical doctor's perspectives in Portugal. *PLoS One.* (2023) 18:e0290613. doi: 10.1371/journal.pone.0290613
- Dalawi I, Isa MR, Chen XW, Azhar ZI, Aimran N. Development of the Malay language of understanding, attitude, practice and health literacy questionnaire on COVID-19 (MUAPHQ C-19): content validity & face validity analysis. *BMC Public Health.* (2023) 23:1131. doi: 10.1186/s12889-023-16044-5
- Alanzi T, Alanazi F, Mashhour B, Altalhi R, Alghamdi A, Al Shubbar M, et al. Surveying hematologists' perceptions and readiness to embrace artificial intelligence in diagnosis and treatment decision-making. *Cureus.* (2023) 15:e49462. doi: 10.7759/cureus.49462
- Rodrigues Alessi M, Gomes HA, Lopes de Castro M, Terumy Okamoto C. Performance of ChatGPT in solving questions from the Progress test (Brazilian National Medical Exam): a potential artificial intelligence tool in medical practice. *Cureus.* (2024) 16:e64924. doi: 10.7759/cureus.64924
- Rodrigues Alessi M, Gomes HA, Oliveira G, Lopes de Castro M, Grenteski F, Miyashiro L, et al. Comparative performance of medical students, ChatGPT-3.5 and ChatGPT-4.0 in answering questions from a Brazilian National Medical Exam: cross-sectional questionnaire study. *JMIR AI.* (2025) 4:e66552. doi: 10.2196/66552
- Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digital Health.* (2023) 2:e0000198. doi: 10.1371/journal.pdig.0000198
- European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Available online at: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (Accessed June 27, 2025).
- World Wide Web Consortium (W3C). (2025). Web Content Accessibility Guidelines (WCAG) 2.1. Available online at: <https://www.w3.org/TR/WCAG21/> (Accessed June 27, 2025).
- Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ.* (1997) 314:572. doi: 10.1136/bmj.314.7080.572
- Foster RC. KR20 and KR21 for some nondichotomous data (it's not just Cronbach's alpha). *Educ Psychol Meas.* (2021) 81:1172–202. doi: 10.1177/0013164421992535
- Mokkink L.B., Elsmann E., Terwee C.B. (2024). COSMIN manual for systematic reviews of outcome measurement instruments. Version 2.0. Amsterdam: Amsterdam UMC. Available online at: https://www.cosmin.nl/wp-content/uploads/COSMIN-manual-V2_final.pdf (Accessed August 06, 2025).

42. LimeSurvey. (2025). The online survey tool - open source surveys. Available online at: <https://www.limesurvey.org> (Accessed June 27, 2025).
43. Wikipedia. (2025). LimeSurvey. Available online at: <https://en.wikipedia.org/wiki/LimeSurvey> (Accessed June 27, 2025).
44. Torvalds L. (2025). Linux Kernel Archives. Available online at: <https://www.kernel.org/> (Accessed June 27, 2025).
45. Apache Software Foundation. (2023). Apache HTTP server 2.4 documentation. Available online at: <https://httpd.apache.org/docs/> (Accessed June 27, 2025).
46. Oracle Corporation. (2025). MySQL 8.0 community edition. Available online at: <https://dev.mysql.com/doc/refman/8.0/en/> (Accessed June 27, 2025).
47. The PHP (2025). Group. PHP Manual. Available online at: <https://www.php.net/manual/en/> (Accessed June 27, 2025).
48. Rescorla E. (2018). The transport layer security (TLS) protocol version 1.3. Internet engineering task force (IETF). Available online at: <https://www.rfc-editor.org/rfc/rfc8446> (Accessed June 27, 2025).
49. Freier A., Karlton P., Kocher P. (2011). The secure sockets layer (SSL) protocol version 3.0. Internet engineering task force (IETF). <https://www.rfc-editor.org/rfc/rfc6101> (Accessed June 27, 2025).
50. Shafranovich Y. (2005). Common format and MIME type for comma-separated values (CSV) files. Internet Engineering Task Force (IETF). Available online at: <https://www.rfc-editor.org/rfc/rfc4180> (Accessed June 27, 2025).
51. Microsoft Corporation (2023). Excel open XML spreadsheet format (XLSX). Available online at: <https://learn.microsoft.com/en-us/openspecs/> (Accessed June 27, 2025).
52. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med.* (2019) 25:30–6. doi: 10.1038/s41591-018-0307-0
53. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. *N Engl J Med.* (2020) 382:1433–6. doi: 10.1056/NEJMp1714229
54. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA.* (2018) 320:2199–200. doi: 10.1001/jama.2018.17163
55. Chen IY, Szolovits P, Ghassemi M. Can AI help reduce disparities in general medical and mental health care? *AMA J Ethics.* (2021) 23:E117–27. doi: 10.1001/amajethics.2019.167
56. Wartman SA, Combs CD. Medical education mousast move from the information age to the age of artificial intelligence. *Acad Med.* (2018) 93:1107–9. doi: 10.1097/ACM.0000000000002044
57. Paranjape K, Schinkel M, Panday RN, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. *JMIR Med Educ.* (2019) 5:e16048. doi: 10.2196/16048
58. Mu Y, He D. The potential applications and challenges of ChatGPT in the medical field. *Int J Gen Med.* (2024) 17:817–26. doi: 10.2147/IJGM.S456659
59. Borenstein J, Howard A. Emerging challenges in AI and the need for AI ethics education. *AI Ethics.* (2021) 1:61–5. doi: 10.1007/s43681-020-00002-7