Developing a Canadian artificial intelligence medical curriculum using a Delphi study

Check for updates
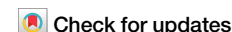
Rohit Singla[1,2,6] ✉, Nikola Pupic[3,6], Seyed-Aryan Ghaffarizadeh[3,6], Caroline Kim[3], Ricky Hu[4], Bruce B. Forster[4,5] & Ilker Hacihaliloglu[4,5]

The integration of artificial intelligence (AI) education into medical curricula is critical for preparing future healthcare professionals. This research employed the Delphi method to establish an expert-based AI curriculum for Canadian undergraduate medical students. A panel of 18 experts in health and AI across Canada participated in three rounds of surveys to determine essential AI learning competencies. The study identified key curricular components across ethics, law, theory, application, communication, collaboration, and quality improvement. The findings demonstrate substantial support among medical educators and professionals for the inclusion of comprehensive AI education, with 82 out of 107 curricular competencies being deemed essential to address both clinical and educational priorities. It additionally provides suggestions on methods to integrate these competencies within existing dense medical curricula. The endorsed set of objectives aims to enhance AI literacy and application skills among medical students, equipping them to effectively utilize AI technologies in future healthcare settings.

The integration of artificial intelligence (AI) in healthcare is transformative, enhancing diagnostic accuracy, risk stratification, and treatment efficiency[1,2]. AI applications have proven effective in various clinical domains, such as reducing documentation burdens for clinicians, improving image interpretation in radiology, supporting intraoperative surgical guidance, and facilitating public health risk stratification[2,3].

However, despite these advancements, the adoption of AI curricular elements in medical education has lagged[4]. Few structured programs thoroughly address both the practical and ethical dimensions of AI in medicine[3–10]. With emerging innovations in AI poised to substantially impact medical practice, interest in training current and future physicians about the technology is growing[1,5,11]. Surveys indicate strong backing for AI literacy training, with 81% of UK physicians[4] and 63% of Canadian medical students recognizing its importance[12], while an overwhelming 93.8% of students in Turkey support structured AI education, particularly in the domains of knowledge, applications, and ethics[13]. This gap highlights the need for AI literacy among medical students to ensure they can effectively leverage these tools in their future clinical practice[5,8]. An interdisciplinary approach is required to balance technical details with the humane aspects of patient care[14]. This approach should include understanding, applying, and critically analyzing AI technologies, aligned with educational frameworks such as Bloom's Taxonomy, a hierarchical classification of cognitive skills—knowledge, comprehension, application, analysis, synthesis, and evaluation—that educators use to develop learning objectives and outcomes[3,8,15]. Currently, a core set of AI competencies has not been identified, although numerous guides and suggested topics exist[9,10].

Given the dense nature of medical training, introducing AI concepts must be done thoughtfully to avoid overwhelming students and educators while still providing the necessary competencies[13,16]. An updated undergraduate medical education (UGME) curriculum should integrate AI education with core medical training, ensuring it complements rather than competes with essential medical knowledge[4,13]. Medical competency is the crux of a clinician's skill set and AI must enhance and not distract from that.

In this study, the Delphi method is employed to identify essential curricular elements for the undergraduate medical student, refining expert opinions through iterative rounds to ensure the curriculum is comprehensive and reflective of current and future healthcare needs[17]. This method captures diverse perspectives from clinical, technical, and educational experts, bridging educational gaps[3,8,11]. By addressing what medical students need to know about AI, this curriculum aims to equip the next generation of

[1]MD/PhD Program, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada. [2]School of Biomedical Engineering, University of British Columbia, Vancouver, BC, Canada. [3]MD Undergraduate Program, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada. [4]Department of Medicine, University of British Columbia, Vancouver, BC, Canada. [5]Department of Radiology, University of British Columbia, Vancouver, BC, Canada. [6]These authors contributed equally: Rohit Singla, Nikola Pupic, Seyed-Aryan Ghaffarizadeh. ✉e-mail: rsingla@ece.ubc.ca

healthcare providers with the competencies required to integrate AI into their clinical workflows confidently[5,7,11,13,14,16].

Results

Table 1 presents the demographics of the 18 Canadian subject matter experts who participated in the study. Among the respondents, 10 were male and 8 were female. Regarding educational qualifications, 8 held a Master's

Table 1 | Demographics of Canadian subject matter experts

Demographic	Respondents (no.)
Sex (n = 18)	
Male	10
Female	8
Degree(s) held (n = 30)	
Master's of Science (or equivalent)	6
Master of Public Health	1
Master of Health Informatics	1
Doctor of Philosophy	11
Doctor of Medicine (or equivalent)	10
Clinical speciality (n = 8)	
Cardiology	1
Diagnostic Radiology	3
General Internal Medicine	1
Neurology	1
Psychiatry	1
Urologic Sciences	1
Province of origin (n = 18)	
Alberta	1
British Columbia	10
Newfoundland and Labrador	1
Nova Scotia	2
Ontario	4
Affiliated institution (n = 28)	
Dalhousie University	2
Memorial University of Newfoundland	1
Queen's University	1
University of Alberta	1
University of British Columbia	9
University of Calgary	1
University of Toronto	2
University of Victoria	1
University of Western Ontario	1
Vancouver General Hospital	1
BC Children's Hospital	1
Canadian Institute for Advanced Research	1
Simon Fraser University	1
St. Paul's Hospital	1
Alberta Machine Intelligence Institute	1
Fraser Health Authority	1
London Health Sciences Centre	1
Vancouver Prostate Centre	1

Provinces not represented include Manitoba, New Brunswick, Prince Edward Island, Quebec, and Saskatchewan. Clinical speciality included having completed a postgraduate residency. Fellowships not included. Individual respondents may have multiple affiliated institutions.

degree, 12 held a Ph.D. and 10 had an MD or equivalent degree. Clinical specialities represented included Cardiology, Diagnostic Radiology, General Internal Medicine, Neurology, Psychiatry, and Urologic Sciences.

Table 2 outlines the learning elements selected for inclusion in the AI curriculum. For the ethics theme, 11 elements, such as identifying regulatory issues in data sharing and explaining the importance of data privacy, reached a consensus in the first round. The legal theme included 11 elements, with consensus achieved in rounds one and two, covering topics like data governance, confidentiality, and liability concerns.

Under the theory theme, 29 elements were selected across three rounds. These elements encompass a range of topics, from understanding statistical concepts to differentiating between types of machine learning and evaluating the economic impact of AI in healthcare. For the application theme, 11 elements were agreed upon, with consensus spanning all three rounds. These elements focus on practical skills like data analysis, integrating AI evidence into clinical decision-making, and validating AI models.

The collaboration theme saw all seven proposed elements achieving consensus in the first round. These elements emphasize strategies for maintaining relationships with AI-focused colleagues, shared decision-making, and identifying opportunities for learning about AI in healthcare. The communication theme included seven elements, all reaching consensus in the first round, covering effective communication strategies with both colleagues and patients, empathetic communication skills, and managing disagreements related to AI.

Lastly, the quality improvement theme comprised six elements, all achieving consensus in the first round. These elements focus on evaluating patient feedback, proposing AI improvements, analyzing current AI applications, and applying user-centered design principles to enhance the user experience of AI in healthcare. Figure 1 summarizes the distribution of a number of learning objectives included by the different themes.

Table 3 lists the learning elements excluded from the AI curriculum. In the theory theme, four elements were excluded in the second round, covering computer structure, hardware components, hardware impact on performance, and basic programming concepts. In the application theme, nine elements were excluded. In the second round, elements such as programming AI models, gradient descent, regularization techniques, backpropagation, data transformation with kernels, anomaly detection, code optimization with vectorization, and using TensorFlow were excluded. In the third round, clustering techniques for unsupervised learning were excluded.

Table 4 shows the learning elements that did not reach a consensus. In the legal theme, issues related to intellectual property and copyright management (L12–L14) lacked agreement. In the theory theme, no consensus was achieved on elements related to programming languages, types of deep learning, and deep learning models (T33–T35). In the application theme, elements such as standardizing data, developing and training AI models, dimensionality reduction techniques, using Keras, and hyperparameter tuning (A21–A26) did not reach consensus.

Practicing M.D. vs. Ph.D. Researchers

Comparisons were performed to see if there was any difference in the way experts rated the curricular elements in round one based on their academic background (M.D.s vs. Ph.D.s). Of the 107 curricular elements, only three elements were found to have statistically significant differences between the two groups. Practicing M.D.s more strongly supported the legal element (L6), while researchers more strongly supported the application elements (A1 and A24). All other elements did not reach significance at $p < 0.05$.

Leave-one-institution-out analysis

The findings support that in the first and third rounds, the removal of members affiliated from UBC had a significant influence on the number of included and undecided elements. Similarly, in the second round, removing

Table 2 | Learning elements selected for inclusion, organized by core theme and round in which consensus was achieved

Theme	Element for inclusion	Consensus round	EPA	CanMEDS role
Ethics				
E1	Identify key regulatory issues surrounding data sharing between healthcare institutions, academic institutions, and private organizations.	1	10	Leader, Health Advocate
E2	Analyze the implications of these regulatory issues on data-sharing practices in healthcare.	1	10	Leader, Health Advocate
E3	Apply appropriate response strategies to comply with regulatory requirements related to data sharing between healthcare institutions.	1	5, 10	Leader, Health Advocate, Professional
E4	Explain the importance of data privacy in the context of using artificial intelligence (AI) with healthcare data.	1	10	Communicator, Health Advocate, Scholar
E5	Define equitable AI and explain its importance in promoting fairness and avoiding bias in AI applications.	1	9, 10	Health Advocate
E6	Define and differentiate between the different types of biases that can appear in AI, including algorithmic, data, and user biases.	1	10	Scholar, Professional
E7	Identify real-world examples of each type of bias and their impact on the effectiveness of AI applications.	1	10	Scholar, Health Advocate
E8	Develop strategies to mitigate and prevent the occurrence of biases in AI applications.	1	5, 10	Scholar, Leader
E9	Apply strategies to promote the use of equitable AI and advocate for its implementation.	1	10	Health Advocate, Leader
E10	Define patient rights and the ethical considerations related to using AI in healthcare.	1	9, 10	Health Advocate, Professional
E11	Explain the importance of respecting patient rights when using AI and describe the potential benefits of doing so.	1	9, 10	Health Advocate, Professional
Legal				
L1	Define data governance and explain its importance when working with AI.	1	10	Leader, Scholar
L2	Explain the importance of confidentiality in healthcare data when using AI.	1	9, 10	Professional, Health Advocate
L3	Identify potential risks to data privacy and best practices when using AI, including relevant legal and regulatory requirements.	1	10	Professional, Scholar
L4	Apply appropriate confidentiality measures to ensure the privacy and security of healthcare data when using AI.	1	4, 5, 10	Professional, Scholar
L5	List and explain the various concerns surrounding liability when using AI in healthcare.	1	9, 10	Professional, Health Advocate
L6	Apply strategies to mitigate liability risks associated with the use of AI in healthcare.	1	9, 10	Professional, Leader
L7	Explain the importance of shared decision-making with AI and the physician's role in shared decision-making with AI.	1	5, 9, 10	Communicator, Collaborator
L8	Understand the legal implications of shared decision-making with AI.	1	10	Leader, Professional
L9	List the key issues surrounding the copyright of AI.	2	10	Leader, Scholar
L10	Identify the key components of a data governance framework and how they relate to AI.	2	10	Leader, Scholar
L11	Apply appropriate data governance measures when working with AI.	2	5, 10	Leader, Scholar, Professional
Theory				
T1	Define and differentiate between statistical concepts of accuracy, F1 score, sensitivity, specificity, positive predictive value, negative predictive value, odds ratio, relative risk, positive and negative likelihood ratios.	1	3, 12	Scholar
T2	Interpret and apply these statistical concepts to real-world healthcare scenarios.	1	3, 4, 12	Scholar, Medical Expert
T3	Understand, interpret, and explain the different types of statistics (descriptive vs inferential).	1	3, 4	Scholar
T4	Understand, interpret, and explain the different types of data (numerical vs categorical).	1	3, 4	Scholar
T5	Understand, interpret, use, and explain common terminology used in AI.	1	3, 4, 10	Scholar
T6	Identify the different domains of healthcare where AI has been successfully applied.	1	3, 4, 10	Scholar, Medical Expert
T7	Evaluate the strengths and benefits of using AI in each domain, including improved accuracy, efficiency, and cost-effectiveness.	1	4, 10	Scholar, Health Advocate
T8	Explain how AI has impacted the quality of patient care and the healthcare industry as a whole.	1	10	Scholar, Health Advocate
T9	Identify the limitations and challenges of using AI in different domains of healthcare.	1	3, 10	Scholar, Health Advocate
T10	Predict and anticipate how the workflow of physicians may change with the implementation of AI.	1	4, 10	Scholar, Leader
T11	Identify techniques that will better facilitate the implementation of AI.	1	10	Scholar, Leader
T12	Understand the basic concepts and principles of machine learning.	1	4	Scholar
T13	Identify and differentiate between different types of machine learning, including supervised, unsupervised, and reinforcement learning.	1	4	Scholar

Table 2 (continued) | Learning elements selected for inclusion, organized by core theme and round in which consensus was achieved

Theme	Element for inclusion	Consensus round	EPA	CanMEDS role
T14	Evaluate the strengths and limitations of each type of machine learning and their applications in healthcare.	1	4, 10	Scholar
T15	Identify and differentiate between different types of regression analyses, including linear, logistic, and Poisson regression.	1	4	Scholar
T16	Understand the concept of model selection in machine learning.	1	4	Scholar
T17	Understand the basic concepts and principles of deep learning.	1	4	Scholar
T18	Understand the different applications of deep learning in healthcare, including image analysis, natural language processing, and time series analysis.	1	4	Scholar, Medical Expert
T19	Understand the basic concepts and principles of natural language processing (NLP).	1	4	Scholar
T20	Identify and differentiate between different applications of NLP in healthcare, including clinical documentation, patient communication, and disease surveillance.	1	4, 6, 7, 10, 12	Scholar, Medical Expert
T21	Evaluate the impact of NLP on the quality and efficiency of healthcare processes.	1	4, 6, 7, 10	Scholar, Medical Expert
T22	Evaluate the strengths and limitations of each type of deep learning and their applications in healthcare.	2	4, 10	Scholar
T23	Identify and differentiate between different types of models, including decision trees, random forests, and support vector machines.	2	4, 10	Scholar
T24	Evaluate the strengths and limitations of each type of model and their applications in healthcare.	2	4, 10	Scholar
T25	Develop skills in data preprocessing, feature engineering, model selection, and evaluation.	2	10	Scholar
T26	Apply these skills to solve real-world problems in healthcare using AI tools.	2	4, 10	Scholar, Medical Expert
T27	Evaluate the economic impact of AI adoption in healthcare, including the costs associated with implementation and maintenance.	3	10, 12	Scholar, Health Advocate
T28	Analyze the potential cost savings and revenue generation opportunities associated with using AI in healthcare.	3	10, 12	Scholar, Health Advocate
T29	Define and differentiate between big data and traditional data sets.	3	–	Scholar
Application				
A1	Analyze and interpret data, including AI model input and output, to inform decision-making.	1	4, 5, 8	Medical Expert, Scholar
A2	Integrate evidence from AI models into clinical decision-making practices in healthcare.	1	3, 4, 5	Medical Expert, Scholar
A3	Critically evaluate the integrity, reliability, and applicability of research on AI applications in healthcare.	1	10	Scholar, Professional
A4	Create research questions that are well-designed and specific to AI research.	2	10	Scholar
A5	Collect and manage data effectively for AI research.	2	10	Scholar
A6	Apply principles of data stewardship to ensure the quality and security of AI data.	2	10	Scholar, Professional
A7	Validate AI models using appropriate statistical methods to ensure their accuracy and reliability for research purposes.	2	10	Scholar, Professional
A8	Use different functions and tools to visualize data in order to gain insights from it.	2	10	Scholar
A9	Preprocess data appropriately for AI research by cleaning, transforming, and selecting relevant features.	3	10	Scholar
A10	Evaluate and select appropriate algorithms for specific AI problems, based on their strengths and limitations.	3	3, 4, 5, 8, 10	Scholar, Medical Expert
A11	Execute and interpret error analysis in machine learning and deep learning models.	3	3, 4, 5, 8, 10	Scholar, Medical Expert
Collaboration				
C1	Develop strategies for establishing and maintaining positive relationships with colleagues involved in the AI side of healthcare, such as data scientists.	1	7, 9	Collaborator, Leader
C2	Distinguish between the roles of a physician, other healthcare providers, and data scientists to promote clear communication.	1	10	Collaborator, Communicator
C3	Engage in shared decision-making with colleagues focused on the AI aspect of healthcare to promote patient-centered care.	1	7, 8, 9	Collaborator, Communicator
C4	Reflect on one's own roles and limitations in the context of AI in healthcare, including ethical considerations and potential biases, to promote responsible use of AI tools.	1	10	Professional, Collaborator
C5	Identify opportunities for learning and self-improvement with respect to one's AI abilities, including training programs and online resources, to ensure that one's skills and knowledge remain up-to-date.	1	10	Scholar, Professional
C6	Identify, select, and navigate credible sources to learn about AI in healthcare, including peer-reviewed publications, expert opinion, and government reports, to ensure that one is using accurate and reliable information.	1	10	Scholar, Professional
C7	Explain the importance of patient inclusion when designing AI for healthcare to ensure that AI tools are designed and implemented in a way that reflects the needs and values of the patient population.	1	10	Health Advocate, Communicator

Table 2 (continued) | Learning elements selected for inclusion, organized by core theme and round in which consensus was achieved

Theme	Element for inclusion	Consensus round	EPA	CanMEDS role
Communication				
Cm1	Predict and anticipate how patient interactions may change with the implementation of AI.	1	7	Communicator, Health Advocate
Cm2	Develop effective communication strategies to disseminate AI-related knowledge and research to colleagues in the healthcare industry.	1	7	Communicator, Scholar
Cm3	Develop patient-friendly materials to disseminate AI-related knowledge and research to patients.	1	5, 10, 12	Communicator, Health Advocate
Cm4	Demonstrate empathetic communication skills when discussing the use of AI in patient care, including patient-centered approaches that encourage patient trust and autonomy.	1	4, 5, 6, 7, 9, 10, 12	Communicator, Health Advocate
Cm5	Manage disagreements and emotionally charged conversations related to AI effectively, including techniques for de-escalation and conflict resolution.	1	7	Communicator, Professional
Cm6	Collect and synthesize relevant information from patients and other sources for use in AI analysis.	1	7	Communicator, Scholar
Cm7	Appropriately interpret and document results from AI analyses for use in patient care and other healthcare decision-making processes.	1	6, 7, 10	Communicator, Medical Expert
Quality improvement				
Q1	Evaluate patient feedback to identify areas of improvement for AI in healthcare.	1	10	Health Advocate, Scholar
Q2	Propose solutions to improve the capability of AI in healthcare based on patient feedback and experience.	1	10	Health Advocate, Leader
Q3	Analyze current applications of AI in healthcare to identify areas for improvement.	1	10	Health Advocate, Professional
Q4	Evaluate community health needs and propose solutions using AI to address these needs.	1	10, 12	Health Advocate, Leader
Q5	Integrate patient feedback into the development and implementation of AI in healthcare.	1	10, 12	Health Advocate, Communicator
Q6	Apply principles of user-centered design to improve the user experience of AI in healthcare.	1	10, 12	Health Advocate, Scholar

The Association of Faculties of Medicine of Canada's entrustable professional activities (EPAs) and the CanMEDS roles are mapped for each included element.
A mapping to entrusted professional activities (EPAs) as well as to CanMEDS competency framework roles is included.

UBC shifted results towards a higher exclusion rate, slightly lower inclusion rate, and decrease in the undecided category. Removing other institutions had a comparatively minor impact on inclusion, exclusion, and undecided categories across all three rounds. A breakdown of results is available in Supplementary Table 1.

Discussion

Overall, 77% ($n = 82$), of the AI curricular elements proposed were deemed to be important for medical students to know how to proficiently use AI, and 77% ($n = 63$) of these included elements that reached consensus in the first round. Thematically, non-technical elements quickly achieved consensus for agreement in the first round. This included unanimous agreement for all elements in ethics (11/11), communication (7/7), collaboration (7/7), and quality improvement (6/6). This highlights the necessity for future physicians to understand AI in a capacity that allows them to engage with it in a safe manner to improve care for patients and ensure transparency in the care they provide. Furthermore, these broader themes already exist and are taught in Canadian UGME, representing an avenue for integration rather than curricular replacement.

The technical themes of theory and application were less decisively included, with only 21/36 and 3/26 elements selected for each one, respectively. The elements included focused on the validation of AI and its strengths and limitations, likely guiding future physicians toward the proper and judicious use of AI. One expert emphasized the importance of medical students understanding the limitations of quantitative data, warning that "the high volumetric quantitative data should not be used to devalue the qualitative data, such as doctor–patient communication and relationship." In a similar view, training should include how to critically appraise AI models for appropriate use in clinical scenarios, akin to evaluating randomized controlled trials.

Regarding the application theme, which had the lowest number of included elements in the first round, we postulate that this is due to the increasing complexity and technicality of the knowledge that physicians use daily. One expert emphasized this point, highlighting that the role of the medical student is the delivery of medical knowledge, not programming. Another expert concurred, adding clinicians should not "be responsible for data collection, cleaning, pre-processing, and the AI model training. These responsibilities deviate from the clinicians' responsibility of caring for patients." Programming and deep learning skills suit engineers, while physicians should validate AI and interpret its output. There will likely be a need for certain physicians to take on a larger role with respect to AI innovation and integration, but the vast majority will be using AI in their everyday practice[8,18]. This explains the exclusion of specific data science techniques and undecided legal elements related to intellectual property.

With respect to our analysis looking at the difference in ratings between expert groups (MDs versus PhD), it was evident that there was no overall difference in rating based on academic background, with all but three comparisons being not statistically significant. This similarity may be attributed to a broad consensus on core elements that are important, underling the complementary expertise of both groups. Additionally, by selecting Ph.D. researchers who have exposure to the medical field, we ensured overlapping yet distinct perspectives. This result also points to the importance of opinions from both practicing M.D.s and Ph.D. researchers, suggesting that their combined insights can lead to a more comprehensive and balanced curriculum.

The leave-one-institution-out showed a difference between included, excluded, and undecided elements when UBC was included versus excluded, highlighting the effect of an increased number of experts from one institution. However, due to the limited sample size, all experts were
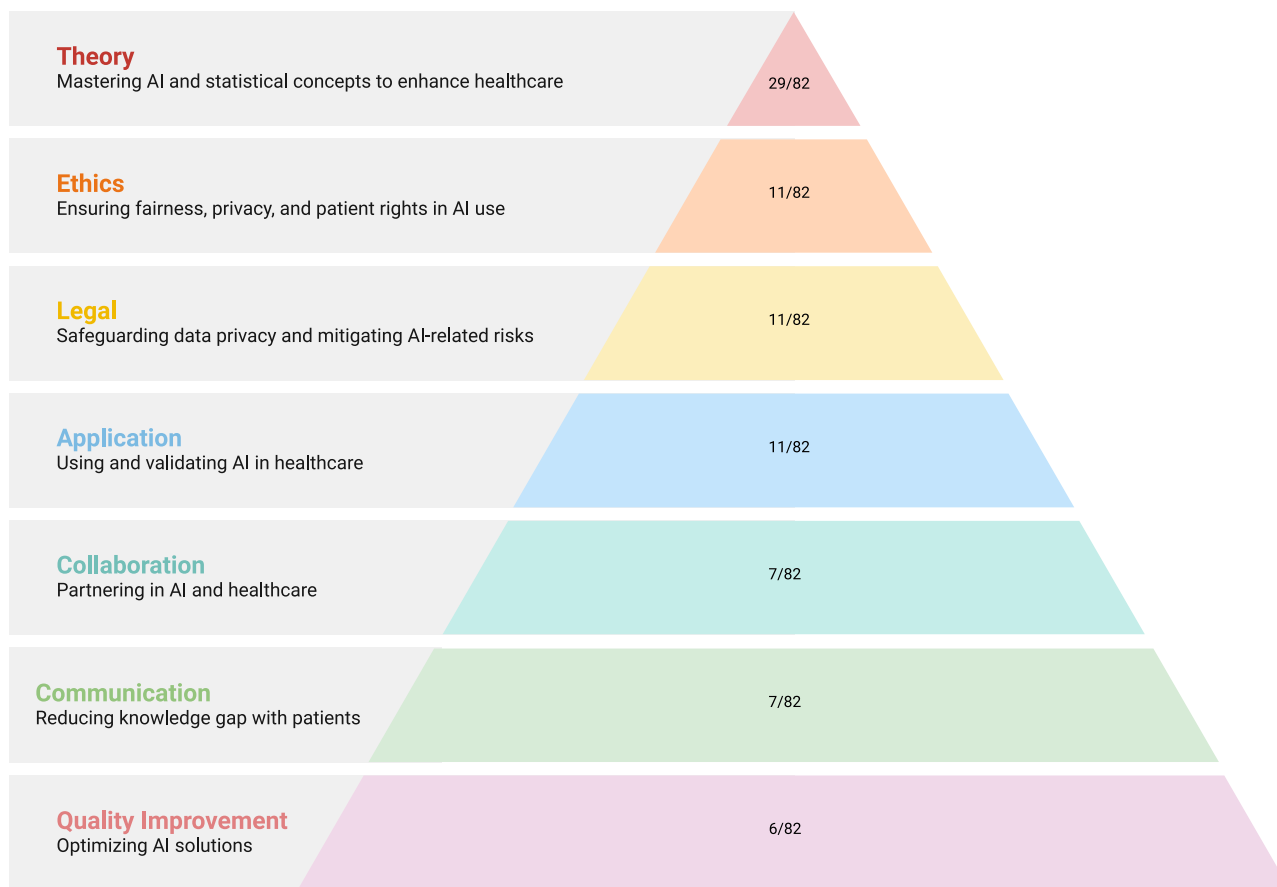
Fig. 1 | Distribution of items by thematic priority in AI healthcare consensus. The figure ranks key themes based on expert consensus, reflecting the proportional focus on each theme relative to the total number of items agreed upon by the experts. The pyramid visually emphasizes the hierarchy of thematic priorities, with 'Theory' at the apex indicating the highest focus, while 'Quality Improvement' occupies the broadest base, suggesting a foundational yet less prioritized theme.

Table 3 | Learning elements selected for exclusion, organized by core theme and round in which consensus was achieved

Theme	Element for exclusion	Consensus round
Theory		
T29	Explain the basic structure and function of a computer, including the central processing unit, memory, and storage.	2
T30	Identify the different types of hardware components and their roles in computer operation.	2
T31	Evaluate the impact of hardware specifications on computer performance and application capabilities.	2
T32	Understand the fundamental concepts of programming, including data types, control structures, functions, and algorithms.	2
Application		
A12	Apply programming concepts to build AI models, tools, and simple healthcare applications.	2
A13	Define gradient descent in machine learning models.	2
A14	Implement regularization techniques to reduce overfitting in models.	2
A15	Understand and apply backpropagation for deep learning models.	2
A16	Use kernels to transform data in machine learning and deep learning models.	2
A17	Understand and apply clustering techniques for unsupervised learning.	3
A18	Implement anomaly detection techniques for identifying outliers in data.	2
A19	Apply vectorization techniques to optimize code in machine learning and deep learning models using Python.	2
A20	Use TensorFlow to build and train deep learning models.	2

included, which may have impacted the overall generalizability of our curricular elements. This is further discussed in the paragraph on limitations.

Although there are no formal existing AI curricula for UGME, there have been efforts to supplement AI education for medical students and residents outside of the curriculum. Lindqwister et al. presented an AI curriculum for radiology residents with didactic sessions and journal clubs, aligning with our elements on AI strengths and limitations (T9) and regulatory issues (E1) to ensure a balanced technical and ethical education[19]. Hu et al. implemented an AI training curriculum for Canadian medical

Table 4 | Learning elements for which no consensus was achieved organized by core theme

Theme	Elements not reaching consensus	% Consensus to include or exclude
Legal		
L12	List the key issues surrounding the intellectual property of AI.	54% in favor of inclusion.
L13	Analyze the implications of intellectual property issues related to the use of AI in healthcare.	69% in favor of inclusion.
L14	Apply appropriate strategies to protect and manage copyright issues when using AI in healthcare.	54% in favor of inclusion.
Theory		
T33	Develop skills in programming languages commonly used in healthcare, such as Python and R.	62% in favor of exclusion.
T34	Identify and differentiate between different types of deep learning, including convolutional neural networks and recurrent neural networks.	54% in favor of inclusion.
T35	Identify and differentiate between different types of models in deep learning, including autoencoders and generative adversarial networks.	54% in favor of exclusion.
Application		
A21	Standardize data to ensure consistency and comparability for AI research purposes.	69% in favor of inclusion.
A22	Develop and implement AI models for research purposes.	54% in favor of exclusion.
A23	Train AI models using appropriate techniques and algorithms, and fine-tune them as needed.	69% in favor of inclusion.
A24	Perform dimensionality reduction techniques such as PCA for feature selection and visualization.	62% in favor of exclusion.
A25	Use Keras to build and train deep learning models.	69% in favor of exclusion.
A26	Perform hyperparameter tuning to optimize model performance.	62% in favor of exclusion.

undergraduates focusing on workshops and project feedback, aligning with our inclusion of elements like applying AI models to clinical decision-making (A2) and developing strategies to mitigate biases (E8)[9]. Krive et al. Created a modular 4-week AI elective for fourth-year medical students, primarily delivered online, aligning with our elements of critical appraisal of AI research/technology (A7, A13), clinical interpretation of results from AI tools (A11), developing strategies to mitigate bias (E8), and communicating results to patients (COM3)[20]. As such, our study builds a framework for medical educators and future research. The UGME curriculum prepares students for generalist practice, covering physiology, anatomy, pathology, diagnostics, therapeutics, clinical decision-making, consultations, and counseling. There is little room for a drastic overhaul of UGME. The University of Toronto's UGME introduces fundamental AI concepts, discussing machine learning, AI's role in healthcare, potential applications, and ethical challenges; showing how AI education can be integrated into UGME, emphasizing core AI literacy and relevance in medicine. Our study findings help identify that nuanced view on which elements should and should not be taught as agreed upon by experts.

The Delphi method, which relies on expert opinion, provided a robust and iterative framework that allowed us to tailor the curriculum to these specific needs, ensuring it is both comprehensive and practical for medical students[17]. We also based our approach on similar studies that have successfully used expert opinions to create or update curricula in medicine for different subject areas, leveraging their structures to ensure our process was thorough[21–25].

In examining the UGME structure, there are several ways to include AI education without significantly impacting the existing curriculum. One approach is to incorporate AI literature into the current biostatistics curriculum, ensuring that students learn to critically appraise and validate AI literature and tools. This integration would also expose students to AI topics and new technologies. Additionally, incorporating AI into facilitator-led case-based learning (CBL) and problem-based learning (PBL) sessions would allow students to explore various AI tools and their impacts[26–30]. These sessions could also provide opportunities to discuss AI ethics topics, such as AI scribes, AI in clinical decision-making, AI policy, and novel AI research. For example, the framework for responsible healthcare machine learning could be discussed in these small groups to explore a simulated process from problem formulation to envisioned deployment[14]. Furthermore, providing hands-on sessions with AI tools currently used in the medical field during clinical rotations, such as point-of-care ultrasound guidance using AI[31] or digital scribes for documentation[32], can help students

improve their technical skills and understand the benefits and risks of these tools. Inviting guest lecturers involved in AI and medicine to discuss the salient principles of AI that medical students need to know and current research in AI and medicine would further enrich their learning experience. Introducing annual modules on AI ethics or baseline knowledge, similar to those required for other rotations, would ensure that students remain up-to-date with the evolving field of AI. Encouraging students to engage in at least one AI-related research project during their 4 years of medical school would deepen their understanding of the subject matter. Additionally, it is important to acknowledge that the integration of an AI curriculum should be flexible and may need to be adapted to fit the specific educational frameworks and resources available at different institutions.

Each included element has been mapped to the AFMC's EPA and CanMEDS roles to underscore the importance of medical AI education. The elements span nearly all EPAs and CanMEDS roles, demonstrating that AI knowledge meets several exit competencies and could be reasonably justified for integration. Alternatively, as existing competencies are updated, specific competencies with regard to select inclusion elements could be included. Endorsement of AI education by a national governing body, supported by a standardized AI curriculum, would encourage medical schools across Canada to integrate AI education into UGME curricula, enhancing future healthcare practitioners' knowledge. As an initial set of suggestions, we mapped each learning objective to potential implementation strategies in Supplementary Table 2.

Our study faced several limitations. Selection bias was a concern using non-probability purposive sampling; despite our efforts to include a diverse and representative group of 106 individuals from across Canada, non-participation could still lead to bias. We recognize that this method can lead to systemic bias and possible over-representation of certain geographic regions or institutes with more established AI programs, thus skewing the curricular elements towards their perspectives. This was seen in our study, with the largest number of experts being from UBC (likely due to the study being conducted by UBC) and a lower response rate from the Atlantic region, Quebec, and the Prairies. The number of experts from UBC resulted in a significant influence on the number of included, excluded, and undecided elements throughout the rounds. These limitations could have been addressed by using random sampling techniques to ensure a more representative sample. Additionally, expanding the pool of experts to a broader range of geographic locations, nationally and internationally, and institutions that may not have medical schools, but other health-related programs associated with them. A larger sample size would also allow for a better
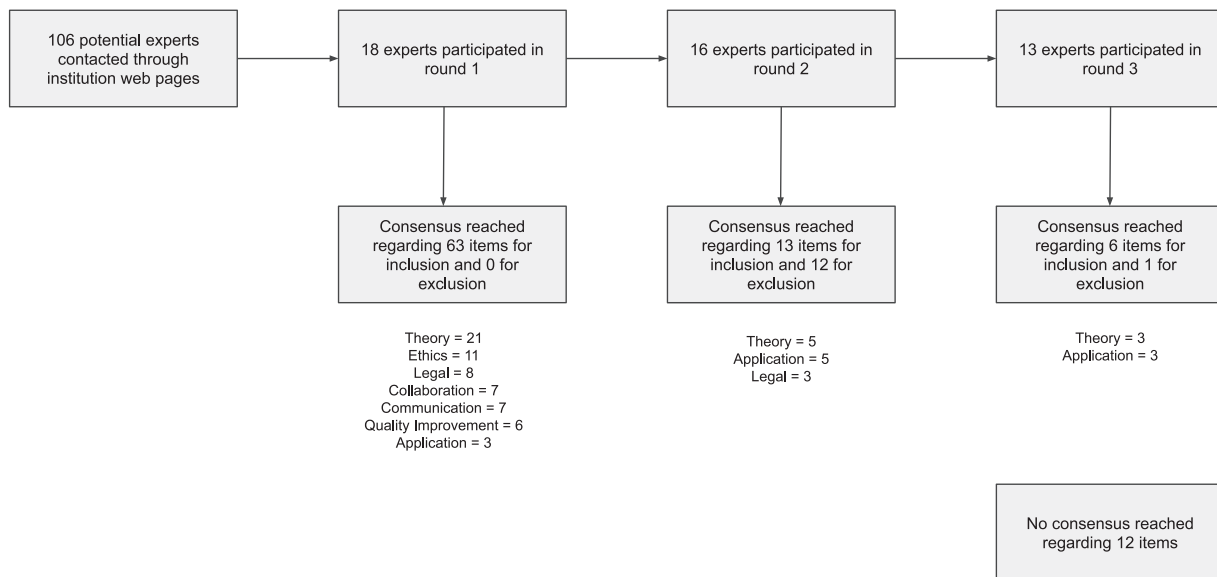
Fig. 2 | This figure illustrates the inclusion and exclusion process of themes in a Delphi study conducted in three rounds. The figure shows the progression and decision points for themes across the three rounds, highlighting the iterative nature of the Delphi method and the involvement of participants at each stage.

investigation into the impact of heterogeneity of responses across centers, as seen in Supplementary Table 1. The small sample size of 18 respondents means that perhaps not all desired perspectives were included. The reasons for dropout at each stay were not specifically elicited and may include time constraints, lack of engagement, competing priorities, or insufficient interest. Our expert inclusion criteria were restricted to M.D.s and Ph.D.s, which may have further restricted the perspectives considered. Broadening the inclusion criteria to include industry and non-university-affiliated experts with relevant AI and medical education expertise could help mitigate this.

Methods
Curricular element selection
Our team's prior thematic analysis of a systematic review demonstrated six key principles for successful AI implementation in medical school curricula using elements compiled from all included studies[8]. These principles are categorized under ethics, theory and application, communication, collaboration, quality improvement, and perception and attitude. Briefly, (a) ethics emphasizes data sharing regulations, privacy, and equity while respecting patient rights; (b) theory and (c) application cover technical skills from statistics to advanced machine learning; (d) communication aims to facilitate understanding of AI tools among healthcare professionals and patients; (e) collaboration highlights multidisciplinary teamwork for shared decision-making and continuous learning in AI use; and (f) quality improvement involves continuous analysis and adaptation of AI tools[6,33–37]. Each of these principles contributes to a holistic AI curriculum, aiming to develop physicians who can thoughtfully integrate AI into patient care. To create the curricular elements, two reviewers (AG and NP) used the previously compiled elements from the thematic analysis to design 107 new curricular elements using Bloom's Taxonomy. These curricular elements were validated by a third reviewer (RS) before consensus was reached through iterative discussions and validation against existing literature (Supplementary Table 2). A mapping of each learning objective to the originally included study is included in Supplementary Table 2. In addition, experts also provided suggested additions or concerns via open-text responses at the end of each section.

Subject matter expert panel selection
In this study, we used purposive sampling to select 106 experts from a variety of fields relevant to medical education and AI[38]. Three authors (A.G., C.K.,

N.P.) created a list of potential experts from across Canada, identified through major universities with accredited medical schools (www.afmc.ca/about/#faculties). A top-down approach was taken to look through each university's faculties of medicine, science, and engineering, followed by departmental and faculty lists, as well as associated speciality societies, research clusters, and special interest groups. Special attention was given to selecting experts from different universities, research clusters, and interest groups to avoid over-representation. Once created, an additional author (RS) was involved in iterating through the list to determine which experts met the inclusion criteria. Experts were identified using current, publicly available information, including but not limited to faculty web pages and departmental profiles.

The panel comprised professionals from diverse geographical regions and healthcare systems, with qualifications from medical doctorates (M.D.) to doctorates in AI-related fields (Ph.D.). Our purposive sampling aimed for diversity in education, medical speciality, geographic location, and institution type. The geographic disparity seen reflects the complex demands of needing local medical AI expertise, which not all schools have.

The inclusion criteria for selecting experts were stringent and multifaceted, requiring candidates to meet at least four out of seven specific conditions: (a) holding an M.D. or a Ph.D. in a relevant field and working in healthcare or computer science, (b) being involved in academic medicine, (c) participating in medical curriculum development, (d) having expertise in medical education, (e) demonstrating leadership in AI and healthcare, (f) producing significant publications in the field, or (g) managing large research departments in AI and healthcare.

All chosen experts provided informed consent for their participation in the study, which was approved by the University of British Columbia (UBC) Research Ethics Board (H22-01695).

Consensus process
The Delphi study aimed to identify the key AI curriculum components for medical education, employing a structured consensus process and multiple rounds of questionnaires (19). The initial survey incorporated all 107 elements and was administered via Qualtrics (Seattle, WA, USA), ensuring anonymity and iterative feedback.

As outlined in Fig. 2, between October 2023 and May 2024, 106 experts were invited by email to participate. In the first round, they provided anonymized demographic data and rated curriculum elements on a 5-point Likert scale, considering items for inclusion if 70% rated them 4 or 5 and for exclusion if 70% rated them 1 or 2. The 70% threshold was based on previous

Delphi studies used for curriculum development[21–25]. Experts could also suggest any competencies they felt were missing or provide their opinion on the topic, through a comments section.

In the subsequent round, items without 70% consensus, and any new competencies were re-evaluated. Feedback and interrater agreement scores were provided to inform their decisions. A simplified 2-point scale was used. The final round allowed experts to decisively conclude on items still lacking consensus, maintaining the 2-point scale to finalize the curriculum elements. This process ensured a rigorous evaluation guided by expert consensus.

For each element selected for inclusion, was subsequently mapped to one or several of the Association of Faculties of Medicine of Canada's entrustable professional activities (EPAs), representing the responsibilities entrusted to a learner in an unsupervised setting[39]. Furthermore, each element was also mapped to a Canadian Medical Education Directives for Specialists (CanMEDS) role, a framework that details the necessary skills for physicians to meet the needs of their patients (www.royalcollege.ca/en/canmeds/canmeds-framework).

Data and statistical analysis

In round one, the responses for each curricular element on the 5-point Likert scale were averaged, and the percentage agreement was calculated by taking the maximal sum of votes that either favored exclusion (rated as 1 or 2), that was indifferent (rated as 3), or that favored inclusion (rated as 4 or 5), and dividing it by the total number of respondents. In rounds two and three, the responses for each remaining curricular element on the 2-point Likert scale were tallied. Percentage agreement was calculated by taking the maximum value and dividing it by the total number of respondents. The percentage agreement was returned to the experts in subsequent rounds to inform them of the expert agreement on items that did not reach a consensus.

Normality and homogeneity tests were performed for each curricular element. The appropriate statistical test, either an unpaired t-test or a Mann–Whitney test, was performed to compare the average ratings between MD and non-MD experts. A leave-one-institution-out analysis was performed to identify if there was a significant difference between the inclusion, exclusion, and undecided rates in each round.

Data availability

All data generated or analyzed during this study are included in this published article.

References

1. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* **25**, 44–56 (2019).
2. Jiang, F. et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc. Neurol.* **2**, 230 (2017).
3. Karalis, V. D. The integration of artificial intelligence into clinical practice. *Appl. Biosci.* **3**, 14–44 (2024).
4. Banerjee, M. et al. The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. *BMC Med. Educ.* **21**, 429 (2021).
5. Pinto dos Santos, D. et al. Medical students' attitude towards artificial intelligence: a multicentre survey. *Eur. Radiol.* **29**, 1640–1646 (2019).
6. Arbelaez Ossa, L. et al. Integrating ethics in AI development: a qualitative study. *BMC Med. Ethics* **25**, 10 (2024).
7. Alowais, S. A. et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med. Educ.* **23**, 689 (2023).
8. Pupic, N. et al. An evidence-based approach to artificial intelligence education for medical students: a systematic review. *PLoS Digit. Health* **2**, e0000255 (2023).
9. Hu, R. et al. Insights from teaching artificial intelligence to medical students in Canada. *Commun. Med.* **2**, 63 (2022).
10. Esteva, A. et al. A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
11. McCoy, L. G. et al. What do medical students actually need to know about artificial intelligence? *Npj Digit. Med.* **3**, 86 (2020).
12. Teng, M. et al. Health care students' perspectives on artificial intelligence: countrywide survey in Canada. *JMIR Med. Educ.* **8**, e33390 (2022).
13. Civaner, M. M., Uncu, Y., Bulut, F., Chalil, E. G. & Tatli, A. Artificial intelligence in medical education: a cross-sectional needs assessment. *BMC Med. Educ.* **22**, 772 (2022).
14. Wiens, J. et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat. Med.* **25**, 1337–1340 (2019).
15. Hwang, K., Challagundla, S., Alomair, M. M., Chen, L. K. & Choa, F.-S. Towards AI-assisted multiple choice question generation and quality evaluation at scale: aligning with bloom's taxonomy. In (eds Denny P et al.) *Workshop: Generative AI for Education (GAIED): Advances, Opportunities, and Challenges,* NeurIPS workshop paper. 1–8 https://gaied.org/neurips2023/files/17/17_paper.pdf (2023).
16. Li, Q. & Qin, Y. AI in medical education: medical student perception, curriculum recommendations and design suggestions. *BMC Med. Educ.* **23**, 852 (2023).
17. Diamond, I. R. et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J. Clin. Epidemiol.* **67**, 401–409 (2014).
18. Xu, Y. et al. Medical education and physician training in the era of artificial intelligence. *Singap. Med. J.* **65**, 159 (2024).
19. Lindqwister, A. L., Hassanpour, S., Lewis, P. J. & Sin, J. M. AI-RADS: an artificial intelligence curriculum for residents. *Acad. Radiol.* **28**, 1810–1816 (2021).
20. Krive, J. et al. Grounded in reality: artificial intelligence in medical education. *JAMIA Open* **6**, ooad037 (2023).
21. Ma, I. W. Y. et al. The Canadian medical student ultrasound curriculum. *J. Ultrasound Med.* **39**, 1279–1287 (2020).
22. Shah, S., McCann, M. & Yu, C. Developing a national competency-based diabetes curriculum in undergraduate medical education: a Delphi Study. *Can. J. Diabetes* **44**, 30–36.e2 (2020).
23. Ellaway, R. H. et al. An undergraduate medical curriculum framework for providing care to transgender and gender diverse patients: a modified Delphi study. *Perspect. Med. Educ.* **11**, 36–44 (2022).
24. Tam, V. C., Ingledew, P.-A., Berry, S., Verma, S. & Giuliani, M. E. Developing Canadian oncology education goals and objectives for medical students: a national modified Delphi study. *CMAJ Open* **4**, E359–E364 (2016).
25. Craig, C. & Posner, G. D. Developing a Canadian curriculum for simulation-based education in obstetrics and gynaecology: a Delphi Study. *J. Obstet. Gynaecol. Can.* **39**, 757–763 (2017).
26. Burgess, A. et al. Scaffolding medical student knowledge and skills: team-based learning (TBL) and case-based learning (CBL). *BMC Med. Educ.* **21**, 238 (2021).
27. Zhao, W. et al. The effectiveness of the combined problem-based learning (PBL) and case-based learning (CBL) teaching method in the clinical practical teaching of thyroid disease. *BMC Med. Educ.* **20**, 381 (2020).
28. Srinivasan, M., Wilkes, M., Stevenson, F., Nguyen, T. & Slavin, S. Comparing problem-based learning with case-based learning: effects of a major curricular shift at two institutions. *Acad. Med.* **82**, 74–82 (2007).
29. McLean, S. F. Case-based learning and its application in medical and health-care fields: a review of worldwide literature. *J. Med. Educ. Curric. Dev.* **3**, JMECD.S20377 (2016).
30. Trullàs, J. C., Blay, C., Sarri, E. & Pujol, R. Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review. *BMC Med. Educ.* **22**, 104 (2022).

31. Wang, H., Uraco, A. M. & Hughes, J. Artificial intelligence application on point-of-care ultrasound. *J. Cardiothorac. Vasc. Anesth.* **35**, 3451–3452 (2021).

32. Coiera, E., Kocaballi, B., Halamka, J. & Laranjo, L. The digital scribe. *Npj Digit. Med.* **1**, 1–5 (2018).

33. Charow, R. et al. Artificial intelligence education programs for health care professionals: scoping review. *JMIR Med. Educ.* **7**, e31043 (2021).

34. Lee, J., Wu, A. S., Li, D. & Kulasegaram, K. M. Artificial intelligence in undergraduate medical education: a scoping review. *Acad. Med.* **96**, S62 (2021).

35. Blease, C. et al. Artificial intelligence and the future of primary care: exploratory qualitative study of UK General Practitioners' views. *J. Med. Internet Res.* **21**, e12802 (2019).

36. Grunhut, J., Wyatt, A. T. & Marques, O. Educating future physicians in artificial intelligence (AI): an integrative review and proposed changes. *J. Med. Educ. Curric. Dev.* **8**, 238212052110368 (2021).

37. Khurana, M. P. et al. Digital health competencies in medical school education: a scoping review and Delphi method study. *BMC Med. Educ.* **22**, 129 (2022).

38. Tongco, M. D. C. Purposive sampling as a tool for informant selection. *Ethnobot. Res. Appl.* **5**, 147–158 (2007).

39. AFMC EPA Working Group. *AFMC Entrustable Professional Activities for the Transition from Medical School to Residency* 1–26 (Education in Primary Care Association, Faculty of Medicine Canada, 2016).

Author contributions
S.-A.G.: Conceptualization, methodology, validation, investigation, data curation, writing—original draft, writing—review and editing, visualization. B.B.F.: Conceptualization, validation, writing— review and editing, supervision, funding acquisition. C.K.: Investigation, data curation, writing—review and editing. I.H.: Conceptualization, validation, writing—review and editing, supervision, funding acquisition. N.P.: Conceptualization, methodology, validation, investigation, data curation, writing—original draft, writing—review and editing, visualization. R.H.: Conceptualization, methodology, investigation, writing—review, and editing. R.S.: Conceptualization, methodology, validation, investigation, writing—original draft, writing—review and editing, project administration. All authors have read and approved the manuscript.

Competing interests
S.-A.G., B.B.F., C.K., R.H., and N.P. each individually declare no financial or non-financial competing interests. I.H. is the co-founder of PONS Incorporated (Newark, NJ, USA) but declares no non-financial competing interests. R.S. has previously acted as a paid consultant for Sonus Microsystems, Bloom Burton, and Amplitude Ventures but declares no non-financial competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-024-01307-1.

Correspondence and requests for materials should be addressed to Rohit Singla.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.