








ORIGINAL RESEARCH ARTICLE

Mapping medical specialty vulnerability to superintelligent AI: A competency-guided generative AI foresight framework

Malik Sallam^{1,2*}, Johan Snbygg^{3,4,5}, Mazin Aljabiri^{5,6}, Edward Cody⁷,
Reem Allateef⁷, Chadia Beaini^{5,8}, and Mohammed Sallam^{5,9,10,11,12}

¹Department of Pathology, Microbiology and Forensic Medicine, School of Medicine, University of Jordan, Amman, Jordan

²Department of Clinical Laboratories and Forensic Medicine, Jordan University Hospital, Amman, Jordan

³Department of Management, Mediclinic City Hospital, Dubai, United Arab Emirates

⁴Department of Anesthesia and Intensive Care, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

⁵Department of General Medicine, College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

⁶Department of Gastroenterology, Mediclinic Parkview Hospital, Dubai, United Arab Emirates

⁷Department of Family Medicine, Mediclinic Parkview Hospital, Dubai, United Arab Emirates

⁸Department of Nephrology/Dialysis, Mediclinic Welcare Hospital, Dubai, United Arab Emirates

⁹Department of Pharmacy, Mediclinic Parkview Hospital, Dubai, United Arab Emirates

¹⁰Department of Management, Mediclinic Parkview Hospital, Dubai, United Arab Emirates

¹¹Department of Management, School of Business, International American University, Los Angeles, California, United States of America

¹²Department of Pharmacology and Therapeutics, College of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

***Corresponding author:**

Malik Sallam
(malik.sallam@ju.edu.jo)

Citation: Sallam M, Snbygg J, Aljabiri M, *et al.* Mapping medical specialty vulnerability to superintelligent AI: A competency-guided generative AI foresight framework. *Artif Intell Health*. 2026;3(2):025340067. doi: 10.36922/AIH025340067

Received: August 20, 2025

Revised: September 9, 2025

Accepted: September 18, 2025

Published online: October 15, 2025

Copyright: © 2025 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract

The evolution of artificial intelligence (AI) raises questions about the future roles of physicians. This study aimed to propose an exploratory foresight model for stratifying risk across medical specialties, using board-defined competencies and generative AI (genAI) evaluation as the assessment tool. We developed a heuristic framework, the Machine automat-ability, Diagnostic Ambiguity, Legal/ethical complexity, Interpersonal intensity, Knowledge codifiability, Evidence in data, Difficulty of procedures (MALIKED) score, to capture dimensions of displacement vulnerability for 27 board-recognized specialties. To minimize individual bias, ratings were generated by three genAI models (ChatGPT, DeepSeek, and Gemini). Data-centric fields—Clinical Pathology (30.3/35), Anatomic/Clinical Pathology (29.3/35), and both Anatomic Pathology and Radiology (28.0/35 each)—clustered in the highest-vulnerability tier. In contrast, procedurally intensive or patient-interaction-heavy specialties—including Psychiatry (11.0/35), Neurosurgery (11.7/35), Obstetrics/Gynecology (13.0/35), General Surgery (13.0/35), Pediatrics (14.3/35), Emergency Medicine (14.3/35), and Family Medicine (14.3/35)—formed the lowest-vulnerability tier. Between these extremes, mixed-mode specialties, such as Internal Medicine (17.0/35) and Neurology (17.0/35), along with Ophthalmology (19.3/35) and Anesthesiology (21.3/35), occupied an intermediate zone. Displacement

risk was driven by knowledge codifiability and data-centricity, while procedural complexity and interpersonal interaction intensity exerted protective effects. This exploratory foresight framework suggests that the risk of displacement by advanced or potentially superintelligent AI is unevenly distributed across medical specialties. While data-driven fields appear most exposed, no specialty is categorically insulated, as multimodal AI and robotics continue to evolve. The MALIKED framework is not predictive but intended as a structured lens for debate, education, and workforce planning regarding the long-term implications of AI in medicine.

Keywords: Generative pre-trained transformer; Clinical competence; Automation; Workforce; Decision support systems

1. Introduction

Debate over artificial intelligence (AI) utility in healthcare has largely emphasized augmentation—the enhancement of physician capabilities rather than their replacement.^{1–3} It has been argued that physicians would remain at the center of healthcare delivery, with AI tools acting as assistants.⁴ The AI-assistance is suggested through automating routine documentation, facilitating administrative work, highlighting relevant guidelines, or flagging potential diagnoses.^{5–8} Consequently, the ultimate responsibility and decision-making authority were expected to remain firmly in human hands, at least through the enforcement of strict governance guidelines on AI use in healthcare.^{5,9} Thus, AI implementation in healthcare was cast as a partner, not a competitor. Alarming, this view is becoming increasingly difficult to sustain in light of the outstanding performance of superintelligent AI models.^{10–12}

The rapid convergence of large language models (LLMs), multimodal foundation models, and robotic process automation is transforming AI tools from narrow-task into broadly capable, adaptive agents.^{13–17} These technologies now integrate vast clinical knowledge, natural language understanding, image interpretation, and even real-time conversational skills within a single interface.^{18,19} In 2024, LLM-based platforms, specifically ChatGPT-4o demonstrated performance on the United States Medical Licensing Examination exceeding 90% across all components, rivaling—and in some cases surpassing—the average board-certified physician.²⁰ In addition, radiology-focused AI tools are reporting cancer detection error rates similar to, or even lower than, those of human readers,^{21,22} and dermatologic AI platforms are delivering diagnostic accuracy on par with specialty-trained clinicians.^{23,24}

The trajectory is clear; as AI capabilities extend beyond discrete sub-tasks toward mastery of entire workflows, the prospect of automation is shifting from a theoretical possibility to an operational reality. This is not merely

a technological milestone; it is a workforce inflection point.^{25,26} If machines can autonomously perform the majority of a medical specialty's core functions—accurately, safely, and at scale—the economic and structural incentives for human displacement become compelling.^{27–29}

However, the impact of this AI-transformation will not be uniform across different medical specialties. Each specialty relies on a distinct combination of cognitive reasoning, procedural dexterity, interpersonal communication, and ethical judgment.^{30–32} Radiologists and pathologists, whose work is predominantly data-centric and interpretable in machine-readable formats, appear highly susceptible to AI transformation and automation.^{33,34} In hybrid specialties, such as internal medicine, emergency medicine, and surgery, vulnerability to AI implementation is expected to vary according to task composition and technological feasibility.^{35–37} However, these remain hypothetical scenarios without hard evidence.

Despite the significance of this issue, the existing literature on AI's potential to displace physicians remains fragmented.^{36,38,39} Most studies are narrowly focused on single specialties, rely on narrative commentary to speculate about automation risk, or appear outdated amid the rapid evolution of AI capabilities.^{40,41} A few surveys have solicited expert opinion or surveyed medical students, but such approaches are inherently subjective and prone to bias, depending on the respondents' personal experience, technological literacy, and professional self-interest.^{42–44} Such approaches provide a valuable perspective but are inherently subjective, influenced by respondents' personal experience, technological literacy, and professional self-interest. While these perspectives are valuable, they lack objectivity and standardization. What remains underdeveloped is a structured, comparative framework that allows medical specialties to be assessed against a common set of criteria, in a manner that is transparent,

replicable, and anchored in formally defined competencies rather than anecdote or perception.

The absence of such a framework has practical consequences. Medical students often face career decisions shaped more by speculation than by systematic evidence; educators lack clear signals regarding which competencies to emphasize for AI resilience; and workforce planners have no standardized means to anticipate substitution pressures across medical specialties. As AI tools advance rapidly, this gap becomes increasingly consequential—not because present models already constitute “superintelligent” systems, but because foresight requires considering scenarios in which AI capabilities substantially exceed today’s boundaries. In this context, the term “superintelligent AI” is employed not to assert its present reality, but as a foresight construct consistent with the technology governance literature, denoting hypothetical systems that outperform humans across broad domains of expertise.^{10,11}

To address this gap, we propose an exploratory framework for stratifying medical specialties’ susceptibility to AI-driven displacement, using board-defined competencies as the foundation for analysis. Every specialty recognized by the American Board of Medical Specialties (ABMS) is governed by certification requirements and training milestones articulated by the Accreditation Council for Graduate Medical Education (ACGME) and equivalent international bodies.^{45–48} These documents, while not free of institutional biases, remain the most authoritative and codified articulation of what each specialty is formally trained and certified to do.^{46,49,50} By grounding our analysis in these sources, we aim to move beyond the variability of individual opinion and toward a systematic, transparent, and reproducible basis for comparison.

The novelty of this study lies in both its analytic foundation and evaluative mechanism. Rather than convening a panel of human experts, we used three independent generative AI (genAI) models to score each specialty against a seven-domain framework: Task automatability, procedural complexity, diagnostic ambiguity, patient interaction intensity, data-centricity, knowledge codifiability (standardization), and ethical/legal complexity. These domains are derived from established literature on automation risk in labor economics, adapted to the clinical context.^{38,51–54} Taken together, they capture both the factors that heighten vulnerability (e.g., data-centricity, codifiability) and those that confer resilience (e.g., procedural difficulty, interpersonal intensity).

This design offers three pragmatic advantages. First, it is replicable in protocol—even if genAI model outputs

evolve—given that both the source documents and scoring rubric are openly accessible. Second, it is comparative, applying the same framework uniformly across all medical specialties, thus enabling direct benchmarking. Third, it is exploratory, providing insight into how advanced genAI models “perceive” specialty vulnerability. Rather than predicting inevitable outcomes, the aim is to stimulate structured debate, identify patterns of risk and resilience, and highlight where empirical follow-up will be most urgent as AI integration into healthcare practice deepens.

2. Data and methods

2.1. Study design

We conducted an exploratory, foresight-oriented risk-mapping exercise to anticipate the relative susceptibility of medical specialties to displacement by superintelligent AI. Rather than treating this as a conventional empirical analysis, the study was designed as a framework-building effort, integrating insights from health workforce automation literature with structured, multidimensional scoring tailored to clinical medicine. The primary objective was to develop a conceptual model capable of highlighting patterns and plausible trajectories, not to provide definitive predictions or statistical certainties. No human subjects or patient data were involved; institutional review board approval was therefore not required.

2.2. Specialty selection and data sources

Specialties were identified using the official ABMS classification, supplemented by subspecialties defined by the ACGME.^{45,55} These taxonomies ensured definitional accuracy and comparability. For analytic clarity, the 27 specialties were further grouped into four broad task-oriented clusters—operative/perioperative, acute/comprehensive care, diagnostic/laboratory, and cognitive/longitudinal—using established competency frameworks and published task analyses.^{30,56–61} This grouping was not intended to be exhaustive, but rather to serve as a heuristic for comparative foresight. The 27 medical specialties were classified by the first and senior authors into four groups according to predominant clinical task profiles, procedural versus cognitive emphasis, data modality, and patient interaction characteristics. This grouping was informed by the ACGME Milestones framework, published medical specialist task analysis, and AI readiness literature (Figure 1).^{47,50,62–89}

High-complexity operative and perioperative specialties encompass fields dominated by invasive interventions, procedural complexity, and perioperative management, where manual dexterity and intraoperative judgment are central. Acute and comprehensive care specialties include disciplines that require broad-spectrum reasoning and rapid

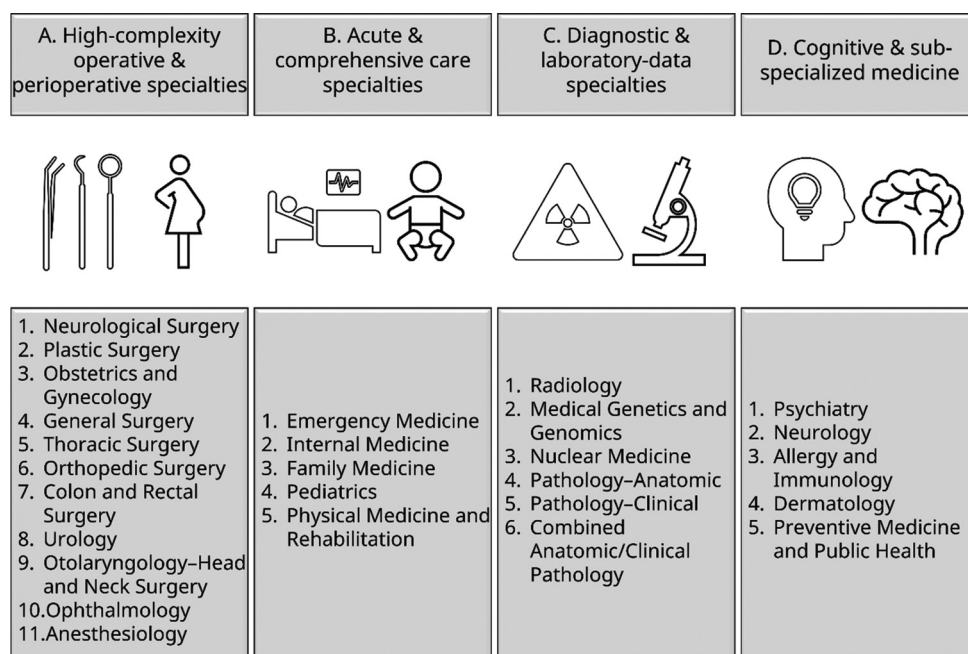


Figure 1. Medical specialty grouping by predominant clinical task profiles

decision-making across organ systems, typically in high-throughput, time-sensitive environments. Diagnostic and laboratory-data specialties are characterized by data-centric workflows, reliance on structured inputs, such as imaging, histopathology, or laboratory assays, and high degrees of standardization. Finally, cognitive and subspecialized medicine includes fields defined by nuanced diagnostic reasoning, longitudinal patient relationships, and ethically complex or highly individualized care contexts.

2.3. Definition of analytical dimensions

We conducted a literature review using Google Scholar through *Publish or Perish* (version 8),⁹⁰ concluding on July 01, 2025. Two independent reviewers (the first and senior authors) undertook thematic coding of all candidate constructs, following constant-comparison principles as highlighted by Naeem *et al.*⁹¹ Codes were iteratively refined, with overlapping constructs consolidated into umbrella categories through consensus adjudication. Constructs were retained only if they were supported by ≥ 2 independent frameworks and demonstrated relevance to the health professions. Discrepancies were resolved through iterative discussion and reference to the most authoritative sources. This process yielded seven consolidated dimensions, balancing parsimony with comprehensiveness (Figure 2). Each of the seven dimensions was assigned equal weight in the analytic model. This decision was based on the exploratory nature of the study and the present absence of validated external benchmarks to justify differential weighting.

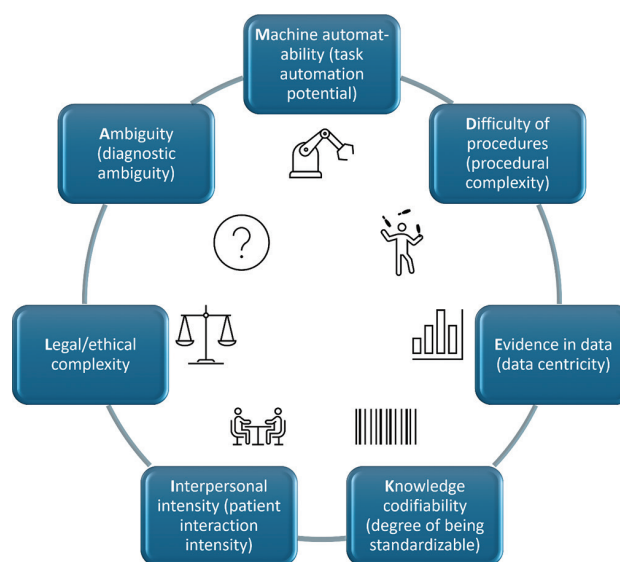


Figure 2. Analytical dimensions of the MALIKED framework for assessing the risk of displacement in medical specialties by superintelligent artificial intelligence. The circular layout represents the equal weighting of all seven dimensions in this study, reflecting a neutral and transparent starting point for analysis.

The seven dimensions were: (i) Task automation potential, defined as the extent to which core medical specialty tasks can be executed by AI given structured data availability, demonstrated algorithmic performance, and prior evidence of automation.^{1,92-97} (ii) Procedural complexity, defined as the dexterous skill, multi-step execution, and intra-procedural adaptability required.⁹⁸⁻¹⁰²

(iii) Diagnostic ambiguity, defined as variability in presentation and the need for nuanced pattern recognition under uncertainty.¹⁰³⁻¹⁰⁶ (iv) Patient interaction intensity, defined as the frequency and centrality of empathic, human-to-human contact.^{30,107-110} (v) Data-centricity, defined as the reliance on large-scale, structured, and digitized data streams, such as imaging or laboratory data.¹¹¹⁻¹¹⁴ (vi) Degree of being standardizable, defined as the uniformity and protocolization of workflows.¹¹⁵⁻¹²⁰ Knowledge codifiability was used synonymously in this context to denote the extent to which tasks are amenable to standardization and formalization into explicit, reproducible processes. (vii) Ethical/legal complexity, defined as the degree of legal liability, ethical deliberation, and professional accountability intrinsic to the specialty.¹²¹⁻¹²⁷

Each analytical dimension was rated on a five-point ordinal scale (1 = minimal, 5 = maximal) by three independent genAI models—GPT-5 (OpenAI), DeepSeek-V3 (DeepSeek Inc.), and Gemini-2.5 Flash (Google)—using a structured prompt (Appendix 1). We deliberately prioritized commercially deployed genAI models over open-weight models, as these are most likely to shape near-term clinical integration. However, we acknowledge that exclusion of open-weight models may limit reproducibility. Gemini Flash was chosen over Gemini Pro for its efficiency and comparable performance on structured tasks. The scoring protocol adhered to the METRICS checklist to standardize the design and reporting of genAI studies in healthcare.¹²⁸ We recognize that genAI judgments are not a substitute for human expert consensus. Rather, they function here as proxy perspectives of frontier AI systems—a foresight-oriented lens into how automation-capable architectures may themselves delineate the evolving boundaries of medical work.

Dimensions were designated as risk-increasing (task automation potential, data-centricity, degree of being standardizable) or risk-mitigating (procedural complexity, diagnostic ambiguity, patient interaction

intensity, ethical/legal complexity). For risk-increasing dimensions, higher scores indicated greater susceptibility to displacement. For risk-mitigating dimensions, reverse scoring was applied: Reversed Score = 5 – Original Score.

Scoring proceeded in five sequential steps. First, GPT-5 (OpenAI), DeepSeek-V3 (DeepSeek Inc.), and Gemini-2.5 Flash (Google) independently generated seven-dimensional score profiles for each specialty based on official board competency descriptions and structured prompts. Prompting concluded on August 10, 2025 (Table 1).

Second, scores were averaged across the three models for each dimension to minimize single-model variance and enhance stability. Third, reverse scoring was applied to the four risk-mitigating dimensions—procedural complexity, diagnostic ambiguity, patient interaction intensity, and ethical/legal complexity—using the transformation, thereby aligning their contribution with overall displacement risk. Fourth, a composite Machine automat-ability, diagnostic Ambiguity, Legal/ethical complexity, Interpersonal intensity, Knowledge codifiability, Evidence in data, Difficulty of procedures (MALIKED) risk score was calculated by summing the three risk-increasing dimensions with the four reversed risk-mitigating dimensions, yielding a theoretical range from 7 (lowest possible risk) to 35 (highest possible risk). Finally, specialties were classified into five risk categories: No risk (7.0), low risk (7.1–14.0), moderate risk (14.1–21.0), high risk (21.1–28.0), and extremely high risk (28.1–35.0). These cutoffs were derived from natural inflection points in the empirical score distribution. We emphasize that these scores are illustrative indices, not empirical measurements. Cutoffs for categorical risk bands (no/low/moderate/high/extremely high) were derived from observed distributional inflection points to aid interpretability, rather than to establish hard thresholds.

2.4. Statistical analysis

All analyses were conducted using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., NY). Because

Table 1. Characteristics of the three generative artificial intelligence (AI) models used in scoring

Characteristic	ChatGPT-5	DeepSeek-V3	Gemini-2.5 Flash
Developer	OpenAI	DeepSeek Inc.	Google DeepMind
Architecture	Proprietary multimodal transformer; integrates text, image, and structured data processing	Transformer-based, optimized for long-context processing	Multimodal transformer with Mixture-of-Experts (MoE) architecture
Strengths	Advanced general-purpose reasoning and contextual interpretation	Capable of handling long-context windows (up to 128K tokens)	High-speed, low-latency inference, optimized for real-time applications and high-throughput tasks. Features an adjustable "thinking budget" to balance performance and cost
Limitations	Closed-source weights; limited transparency in reasoning chains	Lacks native multimodality compared to competitors	Optimized for speed, which may reduce depth or detail in complex, reasoning-heavy tasks

several risk-dimension scores deviated from normality, non-parametric methods were applied. Between-group comparisons of risk-dimension scores across specialties were assessed using the Kruskal–Wallis H test. Inter-rater reliability of genAI-derived scores was evaluated using intraclass correlation coefficients (ICCs), calculated through a two-way random-effects model, enabling the assessment of agreement across clinically coherent specialty clusters. All tests were two-sided, and a $p < 0.05$ was considered statistically significant. These analyses are presented as exploratory diagnostics, intended to assess coherence and contrast across dimensions, not as confirmatory statistical tests.

Hierarchical cluster analysis was applied to classify medical specialties by their multidimensional vulnerability profiles. Ward’s minimum variance method, with squared Euclidean distance, was chosen for its ability to minimize within-cluster variance and yield compact, interpretable groupings. The agglomeration schedule was examined to identify the point at which merging shifted from combining similar specialties to forcing dissimilar ones together. A distinct jump in coefficients occurred between stage 10 (4.167) and stage 11 (5.204), indicating a natural boundary. Accordingly, a rescaled distance cutoff of 5 was selected, balancing internal cohesion with external separation. This threshold was corroborated visually in the dendrogram ensuring clusters were both statistically sound and professionally meaningful.

3. Results

3.1. Overview of exploratory scores across medical specialties

Across the 27 medical specialties evaluated, the MALIKED composite score—an index integrating seven conceptual dimensions of potential susceptibility to AI-driven task displacement—showed wide variation (Table 2). These values should be interpreted not as exact measurements, but as illustrative indices reflecting how genAI models themselves “rate” the relative automability of different specialties when prompted with structured evaluative criteria.

Figure 3 presents the distribution of composite MALIKED scores across specialties, color-coded into heuristic risk bands (green = comparatively lower displacement susceptibility; red = comparatively higher). The purpose of this classification is interpretive visualization—to highlight natural clustering of specialties, rather than to impose validated cutoffs. Composite scores ranged from approximately 11 (Psychiatry) to around 30 (Clinical Pathology). While these numerical differences should not be over-interpreted, the relative spread suggests distinct clusters of susceptibility across the medical profession.

Mean composite scores ranged from 11.00 for Psychiatry, indicating comparatively low automation susceptibility, to

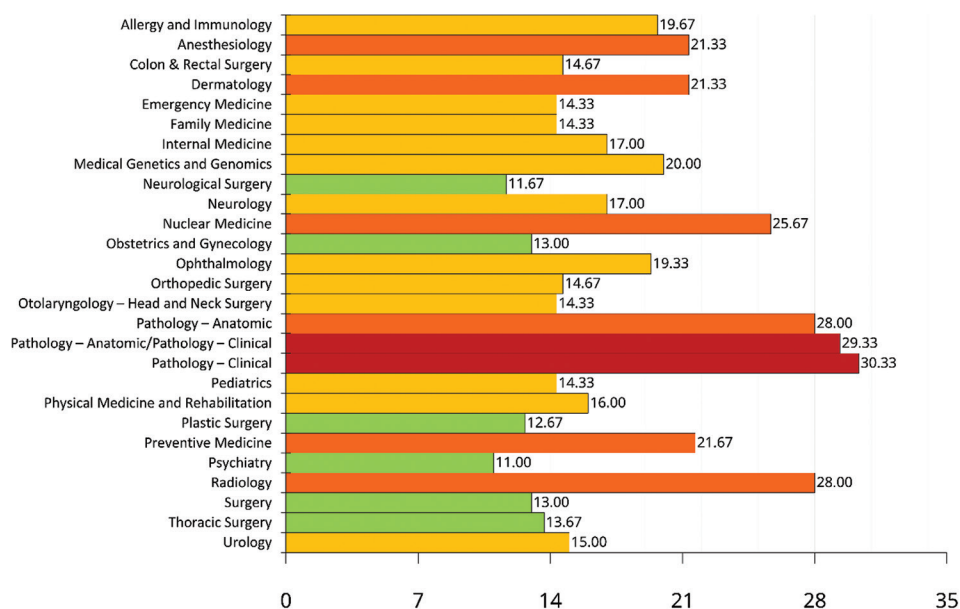


Figure 3. Average MALIKED scores across 27 medical specialties, color-coded by artificial intelligence (AI)-risk category (green = no risk, yellow = low risk, orange = moderate risk, red = high to extremely high risk). Scores integrate seven MALIKED dimensions to quantify susceptibility to displacement by superintelligent AI.

Abbreviation: MALIKED: Machine automat-ability, diagnostic Ambiguity, Legal/ethical complexity, Interpersonal intensity, Knowledge codifiability, Evidence in data, Difficulty of procedures.

Table 2. Risk scores for the seven MALIKED dimensions across 27 medical specialties

Specialty	Task automation potential	Procedural complexity	Diagnostic ambiguity	Patient interaction intensity	Data-centricity	Degree of being standardizable	Ethical/legal complexity
	Mean score out of 5.00	Mean score out of 5.00	Mean score out of 5.00	Mean score out of 5.00	Mean score out of 5.00	Mean score out of 5.00	Mean score out of 5.00
Allergy and Immunology	3.00	3.67	2.00	1.67	3.33	3.33	2.67
Anesthesiology	4.00	1.33	3.00	3.33	4.00	4.33	1.33
Colon and Rectal Surgery	2.33	0.00	2.33	2.67	3.00	3.00	1.33
Dermatology	4.00	2.67	2.67	2.00	3.67	3.67	2.67
Emergency Medicine	3.00	2.00	1.00	1.33	3.33	3.00	0.67
Family Medicine	2.33	4.00	1.00	0.00	2.67	2.33	2.00
Internal Medicine	2.67	3.67	1.00	1.33	3.33	3.00	2.00
Medical Genetics and Genomics	3.33	4.67	1.00	1.67	5.00	3.67	0.67
Neurological Surgery	1.33	0.00	2.00	2.67	3.00	2.33	0.33
Neurology	3.00	3.67	0.67	1.67	3.33	2.67	2.00
Nuclear Medicine	4.00	3.67	2.67	3.67	5.00	4.33	2.33
Obstetrics and Gynecology	2.33	1.00	2.00	1.00	3.00	3.00	0.67
Ophthalmology	3.67	1.00	2.67	2.00	3.67	4.00	2.33
Orthopedic Surgery	2.00	0.00	2.67	3.00	3.00	2.67	1.33
Otolaryngology-Head-and-Neck Surgery	2.33	0.33	2.33	2.00	3.00	3.00	1.33
Pathology-Anatomic	4.67	3.67	3.00	4.67	4.67	4.33	3.00
Pathology-Combined Anatomic/Clinical	5.00	4.00	3.00	4.67	5.00	4.67	3.00
Pathology-Clinical	5.00	4.33	3.33	4.67	5.00	5.00	3.00
Pediatrics	2.33	3.67	1.33	0.00	2.67	2.67	1.67
Physical Medicine and Rehabilitation	2.67	3.00	1.67	1.00	3.00	2.67	2.00
Plastic Surgery	2.00	0.00	2.67	1.33	2.67	2.67	1.33
Preventive Medicine	3.67	4.33	2.00	2.00	4.33	4.00	1.33
Psychiatry	1.67	4.67	0.00	0.00	2.00	1.67	1.00
Radiology	5.00	4.00	2.33	4.00	5.00	5.00	2.67
Surgery	2.00	0.00	2.33	2.67	2.33	2.67	1.00
Thoracic Surgery	2.00	0.00	2.33	3.00	3.00	2.67	0.67
Urology	2.33	1.00	2.33	2.00	3.00	3.00	1.33

Notes: Higher scores indicate greater theoretical susceptibility to displacement by artificial intelligence (AI). Values shown are the mean dimension scores based on the evaluation from three generative AI models (ChatGPT-5, DeepSeek, and Gemini).

30.33 for Clinical Pathology, indicating high automation susceptibility. High scores were also observed in Radiology (28.00) and combined Anatomic/Clinical Pathology (29.33), reflecting high machine automatability, strong data-centricity, and high standardizability. In contrast, specialties with low scores—such as Neurological Surgery (11.67) and Psychiatry (11.00)—were characterized by low procedural standardizability, high patient interaction intensity, and lower automatability potential.

3.2. Internal consistency across AI models

Agreement among the three genAI models was consistently strong (Table 3). ICCs indicated high consistency of

outputs across model architectures, particularly for procedural complexity, patient interaction intensity, and task automation potential. While ICCs do not imply independent human-like reliability, they provide reassurance that results were not idiosyncratic to any single model. Instead, the observed patterns reflect convergent responses across multiple leading systems.

3.3. Variation in MALIKED scores and component domains across specialty groups

When specialties were grouped into four heuristic clusters—operative/perioperative, acute/comprehensive care, diagnostic/laboratory, and cognitive/subspecialized—

systematic differences in displacement susceptibility emerged (Figure 4).

Operative and perioperative specialties consistently showed lower automatability and higher dependence on non-standardized manual skills. Diagnostic and laboratory-data specialties clustered toward higher automatability, standardizability, and data-centricity, aligning with long-standing expectations in the automation literature. Cognitive and subspecialized medicine displayed heterogeneity, with some domains emphasizing ambiguity and interpersonal interaction (e.g., Psychiatry), while others balanced cognitive reasoning with technical procedures (e.g., Neurology). Acute and comprehensive care specialties fell within the middle range, with rapid decision-making requirements offset by partial standardization and moderate data reliance.

3.4. Hierarchical clustering of medical specialties based on genAI scoring

Hierarchical clustering grouped the 27 specialties into four distinct clusters using Ward's minimum variance method and squared Euclidean distance (Figure 5). Cluster 1 included highly digitized, image-driven fields, such as Pathology (Anatomic and Clinical), Radiology, and Nuclear Medicine, reflecting strong alignment with automation potential. Cluster 2 comprised primarily cognitive and diagnostic disciplines, including Internal Medicine, Neurology, Emergency Medicine, and Psychiatry, where human judgment remains central. Cluster 3 included Ophthalmology, Dermatology, and Anesthesiology, specialties that are adopting AI-assisted diagnostic and procedural technologies. Cluster 4 consisted of surgical fields, such as general surgery, orthopedic surgery, and urology, which rely heavily on manual dexterity and

Table 3. Inter-rater reliability of generative artificial intelligence scoring across seven dimensions for 27 medical specialties

Dimension	ICC (single measures)	95% CI (lower-upper)	ICC (average measures)	95% CI (lower-upper)	F (df1, df2)	p-value
Task automation potential	0.886	0.798–0.942	0.959	0.922–0.980	24.305 (26, 52)	<0.001
Procedural complexity	0.904	0.828–0.952	0.966	0.935–0.983	29.250 (26, 52)	<0.001
Diagnostic ambiguity	0.713	0.536–0.845	0.882	0.776–0.942	8.462 (26, 52)	<0.001
Patient interaction intensity	0.902	0.825–0.951	0.965	0.934–0.983	28.746 (26, 52)	<0.001
Data-centricity	0.775	0.624–0.881	0.912	0.833–0.957	11.333 (26, 52)	<0.001
Degree of being standardizable	0.702	0.521–0.838	0.876	0.765–0.939	8.066 (26, 52)	<0.001
Ethical/legal complexity	0.706	0.526–0.840	0.878	0.769–0.940	8.209 (26, 52)	<0.001

Abbreviations: CI: Confidence interval; ICC: Intraclass correlation coefficient.

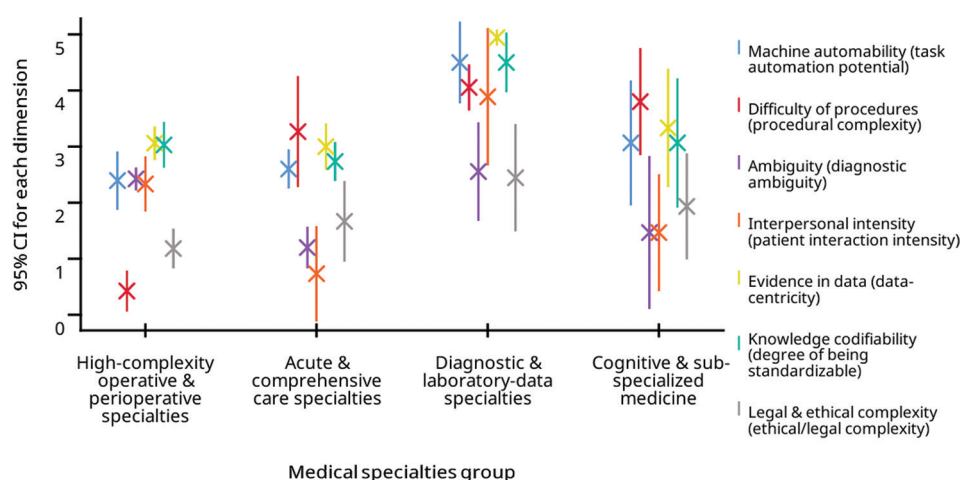


Figure 4. MALIKED domain scores across specialty groups. Mean scores (with 95% confidence intervals) for each of the seven MALIKED dimensions are shown across four specialty clusters: high-complexity operative and perioperative, acute and comprehensive care, diagnostic and laboratory-data, and cognitive and subspecialized medicine.

Abbreviation: MALIKED: Machine automat-ability, diagnostic Ambiguity, Legal/ethical complexity, Interpersonal intensity, Knowledge codifiability, Evidence in data, Difficulty of procedures.

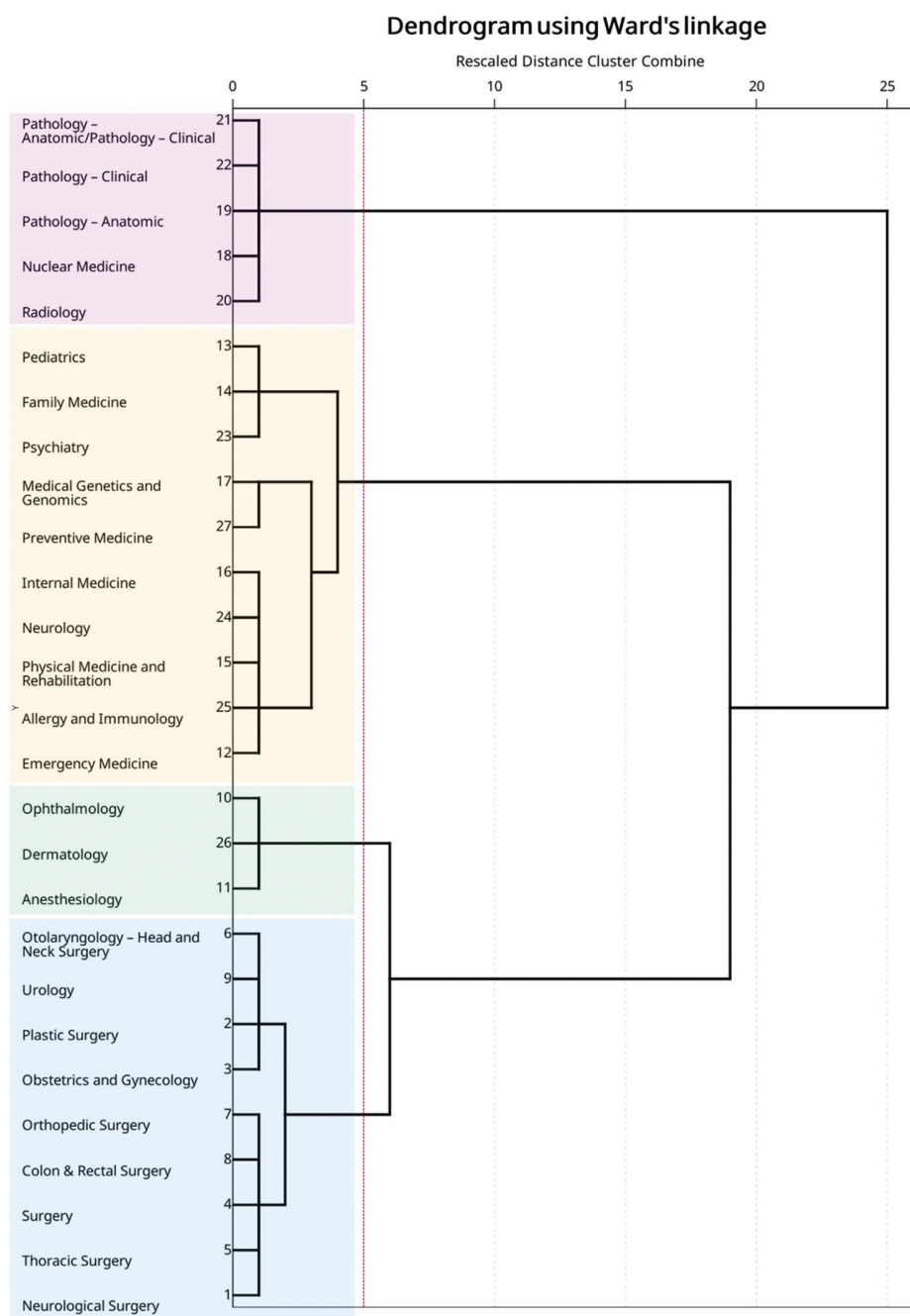


Figure 5. Hierarchical dendrogram of medical specialties based on generative artificial intelligence (genAI) scoring using Ward's minimum variance method with squared Euclidean distance. The red dashed vertical line represents the rescaled distance cutoff of 5, separating the data into four distinct clusters. Each color-coded block corresponds to a cluster, illustrating the natural grouping of specialties according to their AI vulnerability profiles as assessed by three genAI models.

intraoperative decision-making, and are less vulnerable to near-term automation.

4. Discussion

The present study, employing the *ad hoc* MALIKED framework, charts the gradient of vulnerability to

displacement by superintelligent AI across 27 medical and surgical specialties. Although it may seem intuitive that Radiology and Pathology would emerge among the most vulnerable, and Psychiatry among the least, these patterns have, to our knowledge, not been systematically quantified or rigorously compared across the full scope of medical

specialties using a standardized, reproducible analytic framework. By synthesizing seven defining domains—automation potential, procedural complexity, diagnostic ambiguity, intensity of patient interaction, data centrality, codifiability of knowledge, and ethical/legal complexity—the framework constructs a systematic taxonomy of risk.

The spectrum that emerged, ranging from Psychiatry's modest susceptibility (11.00) to Clinical Pathology's heightened exposure (30.33), is consistent with prior expectations and long-standing debates regarding the nature of medical practice itself. Early demonstrations in algorithmic radiology,^{129,130} the pioneering strides of robotic surgery,^{96,131,132} and the steady evolution of digital pathology presaged these divergences,¹³³ though initially, their scope appeared tentative.¹³⁴ In the present era, accelerated by the unprecedented momentum of genAI, these disparities have moved beyond theoretical speculation to manifest as concrete and consequential forces, reshaping not only the configuration of medical practice but also the very competencies upon which the profession has historically defined itself.^{5,10,11,135,136}

In this study, the highest MALIKED scores were observed in the Pathology disciplines—*anatomic, clinical, and their combined variant*—as well as in Radiology. This finding aligns with over a decade of research identifying these specialties as the most susceptible to early AI-driven displacement.^{42,137-139} The convergence of maximal machine automation potential, complete data-centricity, and high knowledge codifiability creates a setting in which algorithmic replication is both technically feasible and economically compelling. Histopathological slides, clinical microbiology and immunology reports, and radiologic images are inherently digital or readily digitized, rendering them especially amenable to convolutional neural networks and transformer-based architectures that now achieve diagnostic accuracies comparable to human experts.^{129,140-144} Much of the clinical laboratory infrastructure is already automated, from complete blood counts and chemistry panels to microbiologic cultures and molecular diagnostics.^{145,146} Recent advances in machine learning extend this trajectory, with models now detecting subtle features in flow cytometry,¹⁴⁷ classifying peripheral blood smears,¹⁴⁸ and even predicting laboratory test utilization errors.¹⁴⁹ These developments underline the feared, accelerating encroachment of superintelligent AI systems into traditionally secure domains of physician expertise.^{43,44,150}

Across medical specialties analyzed in this study, the degree of vulnerability to displacement by superintelligent AI reflects a spectrum—from domains where human presence is indispensable to those where structured

workflows invite algorithmic substitution. At the lowest end of this spectrum lie Psychiatry, Surgery, Pediatrics, and Family Medicine. Psychiatry, with the lowest automation potential and minimal standardizability, exemplifies a discipline irreducibly anchored in conversation, context, and cultural nuance.^{151,152} Digital tools and AI-powered chatbots may assist in triage and monitoring, yet therapeutic alliance, empathy, and embodied presence remain decisive for outcomes.¹⁵³ Pediatrics shares this resilience: Trust-building requires navigating the triangular relationship of child, caregiver, and clinician, while developmental adaptation and parental integration resist algorithmic codification.^{154,155} Family Medicine similarly defies across-the-board automation.^{156,157} Its scope—spanning preventive counseling, longitudinal care, and interpretive flexibility in ambiguous presentations—remains more reliant on relational capital than protocol-driven precision.¹⁵⁸

Moving upward along the gradient, specialties, such as Neurology, Preventive Medicine, Otolaryngology, and Physical Medicine and Rehabilitation occupy a middle ground. Neurology illustrates the duality: While AI excels in seizure detection,¹⁵⁹ magnetic resonance imaging interpretation,^{160,161} and electroencephalogram classification,¹⁶² bedside examination, synthesis across time, and ethically sensitive counseling remain human tasks. Preventive medicine reveals another paradox. Its analytic infrastructure—risk stratification, population surveillance, and enforcement of guidelines—can be largely automated. Yet public health success hinges on persuasion, coalition building, and political negotiation—domains where machines falter.^{163,164} If left unchecked, automation here may hollow the specialty, stripping it of technical functions and leaving practitioners responsible only for the politically fraught. A deliberate redesign of competencies—rewarding leadership, advocacy, and ethical negotiation—could prevent this devaluation. Physical Medicine and Rehabilitation presents a different picture. While gait analysis and outcomes monitoring are easily digitized, the essence of rehabilitation lies in adaptation to human variability, motivational coaching, and psychosocial advocacy.¹⁶⁵ The actionable imperative here is to formalize competencies in narrative coaching and adaptive manual skill as bulwarks against displacement.¹⁶⁶

Procedural specialties, such as urology, thoracic surgery, general surgery, orthopedics, colorectal surgery, plastic surgery, and neurosurgery showed lower susceptibility because their identity is inseparable from dexterity, adaptability, and intraoperative improvisation.^{167,168} Urology highlights this balance: AI algorithms may streamline urinalysis, imaging, and guideline-based management, yet cystoscopy, fertility counseling, and continence surgery require human discretion.¹⁶⁹ Thoracic

and general surgery underline the irreplaceability of improvisation under physiologic instability, bleeding control, and the surgeon's ethical accountability for life-or-death decisions.¹⁷⁰ Orthopedics offers a similar lesson: Robotic planning augments but does not replace the tactile wisdom of fracture reduction or deformity correction.¹⁷¹ Plastic Surgery epitomizes the limits of automatability—esthetic judgment and microsurgical artistry are fundamentally interpretive, not procedural.¹⁷² Obstetrics and Gynecology reflects a duality in which AI readily interprets Pap smears and fetal monitoring, but labor management, surgical emergencies, and ethical counseling demand presence.¹⁷³ Neurosurgery stands at the far end of resilience, where cortical shifts, intraoperative bleeding, and skull-base reconstructions require an improvisational artistry that machines cannot emulate.¹⁷⁴ These domains highlight an urgent imperative: Surgical training must prioritize judgment under uncertainty and ethical accountability—the enduring competencies that safeguard human relevance as technical precision becomes increasingly shared with machines.^{175,176}

Intermediate-vulnerability specialties, such as Internal Medicine, Dermatology, Ophthalmology, Emergency Medicine, Obstetrics and Gynecology, Anesthesiology, and Allergy/Immunology illustrate the spectrum's center. Internal Medicine and Family Practice are increasingly supported by AI copilots managing structured data streams, yet remain indispensable for integrating multimorbidity, reconciling quality-of-life trade-offs, and sustaining longitudinal relationships.^{156,177} Dermatology illustrates the bifurcation between computer vision's triumphs in skin cancer detection and the irreplaceability of surgical artistry and cosmetic ethics.¹⁷⁸ Ophthalmology follows a similar split: Autonomous tools now diagnose diabetic retinopathy, yet microsurgical practice and neuro-ophthalmic reasoning remain firmly human.^{179,180} Emergency Medicine highlights a different kind of resilience: AI can augment triage and early diagnostics, but crisis improvisation and high-stakes procedures resist codification.¹⁸¹ Anesthesiology emerged in this study as more vulnerable than traditionally assumed—a novel finding warranting deeper examination. Increasing standardization of perioperative monitoring, closed-loop anesthesia delivery systems, and robotic assistance in procedural tasks have made portions of anesthesiology highly codifiable and susceptible to automation. Nonetheless, the specialty retains a secure domain in managing airway emergencies, perioperative medical complexity, and rapid, high-judgment decision-making during unexpected deterioration—areas where human oversight remains irreplaceable.¹⁸² Allergy and Immunology likewise straddles automation: Testing and guidelines are easily codified, but rare syndromes,

desensitization, and cultural negotiation require interpretive agility.^{183,184} For this cluster, the actionable step is not merely “augmentation” but redefining training curricula to prioritize human interpretive functions over routine data synthesis, thus ensuring the competencies boards reward remain aligned with what AI cannot replace.

At the highest end of susceptibility lie Radiology, Nuclear Medicine, and Pathology. Radiology has long been the emblematic candidate for automation: Standardized lexicons, digitized images, and repetitive interpretive tasks create near-perfect conditions for machine dominance.^{129,138,185} Routine interpretation is already shifting to algorithmic platforms, with humans retained for edge cases and medicolegal oversight.¹⁸⁶ Nuclear Medicine's vulnerability derives from its inherently digital, protocol-driven workflows, with algorithms now quantifying tumor metabolism, optimizing tracer dosing, and assisting in oncology surveillance.¹⁸⁷ Pathology—both clinical and anatomic—represents the extreme. Much of laboratory medicine is already automated, while digital pathology platforms now rival experts in visual pattern recognition. Here, the actionable imperative is twofold: First, to recast board competencies away from pattern recognition and toward clinical integration, ambiguity resolution, and ethical stewardship; second, to redefine training pipelines so that these specialties evolve into supervisory, consultative, and interpretive roles rather than mere interpretive labor.

Thus, the spectrum from Psychiatry to Pathology illustrates the logic of AI displacement: Specialties grounded in empathy, trust, and artistry remain the most resilient, while those defined by standardized data, repetitive workflows, and high codifiability stand most exposed. The future does not promise uniform substitution, but a reconfiguration in which physicians adapt as supervisors, integrators, and custodians of domains where humanity remains irreducible.

The Kruskal-Wallis test results confirmed that the differences in MALIKED scores were statistically robust across all seven dimensions, underscoring that the gradient of vulnerability to superintelligent AI is not anecdotal but systematic. Procedural complexity and data-centricity emerged as the most discriminating axes, and their significance extends beyond statistical separation. Specialties with high procedural complexity cluster at the low end of automation susceptibility, whereas data-dense, highly codifiable disciplines cluster at the high end. This mirrors the broader trajectory of automation in other knowledge industries: Legal discovery and accounting—fields defined by documents, rules, and standardizable workflows—have already been heavily automated,¹⁸⁸ while craft-like domains—

though technologically augmented—remain resistant to substitution. The ICCs add weight to the framework's validity, with excellent agreement for procedural complexity, patient interaction, and automatability, and good agreement for diagnostic ambiguity, standardizability, and ethical-legal complexity. The particularly strong concordance for procedural complexity suggests that even disparate AI models recognize what practicing surgeons have always known: That dexterity, adaptability, and improvisation at the point of care create a barrier to displacement not easily overcome by machines.

The implications are profound. A uniform “AI-readiness” curriculum would be inadequate, because specialties do not face equivalent risks. Instead, differentiated adaptation strategies are necessary. In high-risk fields, such as radiology, pathology, and nuclear medicine—where data-centricity is maximal and knowledge codifiability high—the human role is shifting from interpretation to oversight, adjudication, and exception handling. Here, medical knowledge must expand beyond mastery of diagnostic content to include literacy in AI validation, algorithmic bias, and integration of machine outputs into complex clinical contexts.¹⁸⁹ Systems-based practice becomes equally central, requiring radiologists and pathologists to oversee not just their immediate diagnostic domain but the broader health system consequences of deploying automated workflows.^{34,129}

In surgical specialties, the principal defense against automation lies not in technical precision but in judgment under uncertainty. Evidence from crisis management shows that intraoperative adaptability, salvage decision-making, and bleeding control remain decisive for outcomes under physiologic instability—domains where tactile intelligence and human accountability are irreplaceable.^{190,191} Competency frameworks should therefore emphasize explicit training and assessment in improvisation, crisis response, and ethical accountability, ensuring that surgeons remain indispensable even as robots handle routine steps. Relational specialties—Psychiatry, Pediatrics, Family Medicine—derive protection less from procedural complexity than from trust, empathy, and cultural negotiation. Psychotherapy research demonstrates that the therapeutic alliance predicts outcomes as strongly as treatment modality,¹⁹² while continuity in Family Medicine drives effective management of multimorbidity.¹⁹³ These specialties may need deliberate workforce expansion as diagnostic fields contract, but such redistribution depends on reimbursement models that reward relational care, rather than leaving labor markets to adjust passively.

Intermediate-vulnerability specialties—including Internal Medicine, Dermatology, Ophthalmology,

Emergency Medicine, Anesthesiology, and Allergy and Immunology—will require hybrid strategies. Their workflows illustrate bifurcation: Algorithmic triage, image analysis, and anesthetic monitoring are partially automated, yet crisis improvisation, rare-case recognition, and ethical discernment remain human. Here, practice-based learning and systems-based practice are paramount. Clinicians must continuously recalibrate scope, delegating routine functions to algorithms while retaining accountability for integration, ambiguity, and patient-centered trade-offs. Boards can operationalize this by incorporating simulation and Objective Structured Clinical Examination-based assessments that explicitly test crisis response and ambiguity management alongside technical proficiency.¹⁹⁴

At the regulatory level, the MALIKED framework highlights asymmetric pressures with clear policy relevance. Diagnostic specialties may contract unless they expand into algorithmic stewardship, population-level data governance, and validation of AI pipelines, while relational specialties may require deliberate growth. Professionalism will acquire new dimensions as physicians are held accountable for supervising machines whose decisions carry clinical consequences.¹⁹⁵ Practice-based learning will become a career-long requirement as competencies evolve alongside technology. Thus, the spectrum from Psychiatry to Pathology functions not merely as a descriptive gradient but as a prescriptive map. It identifies competencies—empathy, improvisation, ethical accountability—that must be codified, assessed, and reimbursed as deliberately as procedural skills. The task is not to resist automation, but to shape it: Positioning physicians as supervisors, integrators, and moral custodians in an era where intelligence itself is no longer uniquely human.

Finally, it is important to recognize that this study has inherent limitations, which should inform a cautious interpretation of the results. The MALIKED framework reduces the complexity of clinical practice to seven dimensions. Although grounded in a comprehensive literature review, the initial step of defining these seven dimensions involved an element of human judgment. Alternative conceptual models might emphasize different constructs, and this potential subjectivity should be acknowledged. Methodologically, we did not validate MALIKED scores against external benchmarks, such as automation adoption rates, residency trends, or policy signals, and deliberately refrained from estimating unemployment probabilities. The framework, therefore, offers an exposure taxonomy rather than a predictive forecast. Contextual variability further constrains interpretation: Risk profiles will shift with advances in robotics, explainable AI, liability rules, and patient

expectations. In addition, ABMS/ACGME-defined competencies do not map neatly onto global health systems. This pilot study also excluded certain specialties and did not capture second-order ripple effects across fields. These limitations underline that MALIKED is best understood as a replicable exploratory protocol—intended to catalyze structured debate and provide a foundation for future empirical validation, not as a deterministic map of medicine's future.

5. Conclusion

This study demonstrated that superintelligent AI pressure will not affect all medical specialties equally. Diagnostic fields, such as Pathology and Radiology—whose workflows are already highly digitized—appear most vulnerable, whereas specialties marked by procedural complexity or intensive human interaction (e.g., Neurosurgery, Psychiatry) show relative resilience. However, no specialty is entirely insulated: Even resilient fields contain automatable sub-tasks, and emerging robotics and multimodal AI may shift present boundaries. The MALIKED framework provides a replicable, competency-based method for systematically assessing these gradients of risk. By anchoring analysis in board-defined competencies and using genAI as a consistent evaluator, the framework offers a structured foundation for anticipating workforce shifts. The implications are clear. Education should reinforce the uniquely human competencies—adaptability, judgment under uncertainty, ethical accountability—that remain difficult to automate. Certification bodies can recalibrate competencies to emphasize oversight and integration, while workforce and policy planning should prepare for asymmetric impacts. Diagnostic fields may potentially be redeployed into supervisory and data-governance roles, while relational fields may require deliberate capacity expansion. The MALIKED framework thus offers a replicable, competency-based tool for systematically assessing AI-driven displacement risk and guiding structured workforce planning.

Acknowledgments

The authors declare the use of generative artificial intelligence (AI) in the preparation of this manuscript. ChatGPT-4o was employed for language refinement (including improvements in grammar, sentence structure, and readability) and technical writing assistance (such as suggesting more effective ways to structure complex technical descriptions). All AI-assisted processes were critically reviewed by the authors to ensure the integrity and reliability of the results. The final decisions and interpretations presented in this article were made solely by the authors.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Malik Sallam

Data curation: All authors

Formal analysis: Malik Sallam, Mohammed Sallam

Investigation: All authors

Methodology: All authors

Project administration: Malik Sallam

Supervision: Malik Sallam, Mohammed Sallam

Validation: Malik Sallam, Mohammed Sallam

Visualization: Malik Sallam

Writing – original draft: Malik Sallam

Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data presented in this study are available from the corresponding author, Malik Sallam, upon reasonable request.

References

1. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: Transforming the practice of medicine. *Future Healthc J.* 2021;8(2):e188-e194.
doi: 10.7861/fhj.2021-0095
2. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: Transforming healthcare in the 21st century. *Bioengineering (Basel).* 2024;11(4):337.
doi: 10.3390/bioengineering11040337
3. Alrassi J, Katsurakis PJ, Chandran L. Technology can augment, but not replace, critical human skills needed for patient care. *Acad Med.* 2021;96(1):37-43.
doi: 10.1097/ACM.0000000000003733
4. Kim J, Chen ML, Rezaei SJ, *et al.* Artificial intelligence tools in supporting healthcare professionals for tailored patient care. *NPJ Digit Med.* 2025;8(1):210.
doi: 10.1038/s41746-025-01604-3
5. Sallam M. ChatGPT utility in healthcare education,

- research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. 2023;11(6):887.
doi: 10.3390/healthcare11060887
6. Patel SB, Lam K. ChatGPT: The future of discharge summaries? *Lancet Digit Health*. 2023;5(3):e107-e108.
doi: 10.1016/s2589-7500(23)00021-3
7. Reddy S. Generative AI in healthcare: An implementation science informed translational path on application, integration and governance. *Implement Sci*. 2024;19(1):27.
doi: 10.1186/s13012-024-01357-9
8. Rabbani SA, El-Tanani M, Sharma S, *et al*. Generative artificial intelligence in healthcare: Applications, implementation challenges, and future directions. *BioMedInformatics*. 2025;5(3):37.
doi: 10.3390/biomedinformatics5030037
9. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*. 2024;10(4):e26297.
doi: 10.1016/j.heliyon.2024.e26297
10. Sallam M, Beaini C, Mijwil MM, Sallam M. Microsoft AI diagnostic orchestrator (MAI-DxO) promises cost savings in diagnosis - but at what cost to the medical profession? *J Stud Med Sci*. 2025;1(1):1515.
doi: 10.20849/jsms.v1i1.1515
11. Sallam M, Beaini C, Mijwil MM, Sallam M. AI at the crossroads of cost and care: A SWOT analysis of microsoft AI diagnostic orchestrator (MAI-DxO) in clinical diagnosis. *J Clin Med Res*. 2025;2(1):1-14.
doi: 10.37155/3060-8708-02
12. Nori H, Daswani M, Kelly C, *et al*. Sequential Diagnosis with Language Models. *arXiv*. Preprint posted online 2025.
doi: 10.48550/arXiv.2506.22405
13. Xiao X, Liu J, Wang Z, *et al*. Robot learning in the era of foundation models: A survey. *Neurocomputing*. 2025;638:129963.
doi: 10.1016/j.neucom.2025.129963
14. Wang J, Shi E, Hu H, *et al*. Large language models for robotics: Opportunities, challenges, and perspectives. *J Automat Intell*. 2025;4(1):52-64.
doi: 10.1016/j.jai.2024.12.003
15. Ye J, Tang H. Multimodal Large Language Models for Medicine: A Comprehensive Survey. *arXiv*. Preprint posted online 2025.
doi: 10.48550/arXiv.2504.21051
16. Kim K, Windle J, Christian M, *et al*. Framework for integrating large language models with a robotic health attendant for adaptive task execution in patient care. *Appl Sci*. 2024;14(21):9922.
doi: 10.3390/app14219922
17. Clusmann J, Kolbinger FR, Muti HS, *et al*. The future landscape of large language models in medicine. *Commun Med (Lond)*. 2023;3(1):141.
doi: 10.1038/s43856-023-00370-1
18. Attrach RA, Moreira P, Fani R, Umeton R, Fiske A, Celi LA. Conversational LLMs Simplify Secure Clinical Data Access, Understanding, and Analysis. *arXiv*. Preprint posted online 2025.
doi: 10.48550/arXiv.2507.01053
19. Singhal K, Azizi S, Tu T, *et al*. Large language models encode clinical knowledge. *Nature*. 2023;620(7972):172-180.
doi: 10.1038/s41586-023-06291-2
20. Bicknell BT, Butler D, Whalen S, *et al*. ChatGPT-4 omni performance in USMLE disciplines and clinical skills: Comparative analysis. *JMIR Med Educ*. 2024;10:e63430.
doi: 10.2196/63430
21. Killock D. AI outperforms radiologists in mammographic screening. *Nat Rev Clin Oncol*. 2020;17(3):134-134.
doi: 10.1038/s41571-020-0329-7
22. Caldas FAA, Caldas HC, Henrique T, *et al*. Evaluating the performance of artificial intelligence and radiologists accuracy in breast cancer detection in screening mammography across breast densities. *Eur J Radiol Artif Intell*. 2025;2:100013.
doi: 10.1016/j.ejrai.2025.100013
23. Pham TC, Luong CM, Hoang VD, Doucet A. AI outperformed every dermatologist in dermoscopic melanoma diagnosis, using an optimized deep-CNN architecture with custom mini-batch logic and loss function. *Sci Rep*. 2021;11(1):17485.
doi: 10.1038/s41598-021-96707-8
24. Salinas MP, Sepúlveda J, Hidalgo L, *et al*. A systematic review and meta-analysis of artificial intelligence versus clinicians for skin cancer diagnosis. *NPJ Digit Med*. 2024;7(1):125.
doi: 10.1038/s41746-024-01103-x
25. Acemoglu D, Restrepo P. Automation and new tasks: How technology displaces and reinstates labor. *J Econ Perspect*. 2019;33:3-30.
doi: 10.1257/jep.33.2.3
26. Qiao D, Rui H, Xiong Q. AI and Jobs: Has the inflection point arrived? Evidence from an Online Labor Platform. *SSRN Electronic Journal*. Preprint; 2023.
doi: 10.2139/ssrn.4656716
27. Liu J, Xu X, Li Y, Tan Y. "Generate: The Future of Work

- through AI: Empirical Evidence from Online Labor Markets. *SSRN Electronic Journal*. Preprint; 2023.
doi: 10.2139/ssrn.4529739
28. Demirci O, Hannane J, Zhu X. Who is AI replacing? The impact of generative AI on online freelancing platforms. *Manag Sci*. 2025;71:8097-8108.
doi: 10.1287/mnsc.2024.05420
29. Teutloff O, Einsiedler J, Kässi O, Braesemann F, Mishkin P, del Rio-Chanona RM. Winners and losers of generative AI: Early evidence of shifts in freelancer demand. *J Econ Behav Organ*. 2025;235:106845.
doi: 10.1016/j.jebo.2024.106845
30. Zelesniack E, Oubaid V, Harendza S. Defining competence profiles of different medical specialties with the requirement-tracking questionnaire - a pilot study to provide a framework for medial students' choice of postgraduate training. *BMC Med Educ*. 2021;21(1):46.
doi: 10.1186/s12909-020-02479-6
31. Mullola S, Hakulinen C, Presseau J, et al. Personality traits and career choices among physicians in Finland: Employment sector, clinical patient contact, specialty and change of specialty. *BMC Med Educ*. 2018;18(1):52.
doi: 10.1186/s12909-018-1155-9
32. Mirvis DM. Choosing a medical specialty: The difference between what students want and what society needs. *Israel J Health Policy Res*. 2013;2(1):18.
doi: 10.1186/2045-4015-2-18
33. Nensa F. The future of radiology: The path towards multimodal AI and superdiagnostics. *Eur J Radiol Artif Intell*. 2025;2:100014.
doi: 10.1016/j.ejrai.2025.100014
34. Gaffney H, Mirza KM. Pathology in the artificial intelligence era: Guiding innovation and implementation to preserve human insight. *Acad Pathol*. 2025;12(1):100166.
doi: 10.1016/j.acpath.2025.100166
35. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The aspects of running artificial intelligence in emergency care; a scoping review. *Arch Acad Emerg Med*. 2023;11(1):e38.
doi: 10.22037/aaem.v11i1.1974
36. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*. 2019;7:e7702.
doi: 10.7717/peerj.7702
37. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Med Educ*. 2023;23(1):689.
doi: 10.1186/s12909-023-04698-z
38. Fogo AB, Kronbichler A, Bajema IM. AI's threat to the medical profession. *JAMA*. 2024;331(6):471-472.
doi: 10.1001/jama.2024.0018
39. Rony MKK, Parvin MR, Wahiduzzaman M, Debnath M, Bala SD, Kayesh I. "I wonder if my years of training and expertise will be devalued by machines": Concerns about the replacement of medical professionals by artificial intelligence. *SAGE Open Nurs*. 2024;10:23779608241245220.
doi: 10.1177/23779608241245220
40. Karches KE. Against the iDoctor: Why artificial intelligence should not replace physician judgment. *Theor Med Bioeth*. 2018;39(2):91-110.
doi: 10.1007/s11017-018-9442-3
41. Shonhe L, Min Q. Mitigating AI-induced professional identity threat and fostering adoption in the workplace. *AI Soc*. 2025;40(5):4079-4092.
doi: 10.1007/s00146-024-02170-0
42. Sarwar S, Dent A, Faust K, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digital Med*. 2019;2(1):28.
doi: 10.1038/s41746-019-0106-0
43. Sallam M, Al-Mahzoum K, Almutairi YM, et al. Anxiety among medical students regarding generative artificial intelligence models: A pilot descriptive study. *Int Med Educ*. 2024;3(4):406-425.
doi: 10.3390/ime3040031
44. Sallam M, Al-Mahzoum K, Alaraji H, et al. Apprehension toward generative artificial intelligence in healthcare: A multinational study among health sciences students. *Front Educ*. 2025;10:1542769.
doi: 10.3389/educ.2025.1542769
45. The American Board of Medical Specialties (ABMS). *Standards for Initial Certification*. 14 August; 2025. Available from: <https://www.abms.org/board-certification/board-certification-standards/standards-for-initial-certification> [Last accessed on 2025 Aug 14].
46. Andolsek KM, Jones MD Jr., Ibrahim H, Edgar L. Introduction to the milestones 2.0: Assessment, implementation, and clinical competency committees supplement. *J Grad Med Educ*. 2021;13(2 Suppl):1-4.
doi: 10.4300/jgme-d-21-00298.1
47. Edgar L, Roberts S, Holmboe E. Milestones 2.0: A step forward. *J Grad Med Educ*. 2018;10(3):367-369.
doi: 10.4300/jgme-d-18-00372.1
48. Day SH, Nasca TJ. ACGME international: The first 10 years. *J Grad Med Educ*. 2019;11(4 Suppl):5-9.
doi: 10.4300/jgme-d-19-00432
49. Ekpenyong A, Padmore JS, Hauer KE. The purpose,

- structure, and process of clinical competency committees: Guidance for members and program directors. *J Grad Med Educ.* 2021;13(2 Suppl):45-50.
doi: 10.4300/jgme-d-20-00841.1
50. Potts JR 3rd. Assessment of competence: The accreditation council for graduate medical education/residency review committee perspective. *Surg Clin North Am.* 2016;96(1):15-24.
doi: 10.1016/j.suc.2015.08.008
51. Filippi E, Bannò M, Trento S. Automation technologies and their impact on employment: A review, synthesis and future research agenda. *Technol Forecast Soc Change.* 2023;191:122448.
doi: 10.1016/j.techfore.2023.122448
52. Cattaneo MA, Gschwendt C, Wolter SC. How scary is the risk of automation? Evidence from a large-scale survey experiment. *J Econ Behav Organ.* 2025;235:107034.
doi: 10.1016/j.jebo.2025.107034
53. Khanna NN, Maindarkar MA, Viswanathan V, et al. Economics of artificial intelligence in healthcare: Diagnosis vs. treatment. *Healthcare (Basel).* 2022;10(12):2493.
doi: 10.3390/healthcare10122493
54. Hendrix N, Veenstra DL, Cheng M, Anderson NC, Verguet S. Assessing the economic value of clinical artificial intelligence: Challenges and opportunities. *Value Health.* 2022;25(3):331-339.
doi: 10.1016/j.jval.2021.08.015
55. The Accreditation Council for Graduate Medical Education (ACGME). *ACGME Home: Specialties.* ACGME; 2025. Available from: <https://www.acgme.org/specialties> [Last accessed on 2025 Aug 14].
56. Englander R, Cameron T, Ballard AJ, Dodge J, Bull J, Aschenbrenner CA. Toward a common taxonomy of competency domains for the health professions and competencies for physicians. *Acad Med.* 2013;88(8):1088-1094.
doi: 10.1097/ACM.0b013e31829a3b2b
57. General Medical Council. *GMC Approved Postgraduate Curricula*; 2025. Available from: <https://www.gmc-uk.org/education/standards-guidance-and-curricula/curricula> [Last accessed on 2025 Aug 18].
58. European Union of Medical Specialists. *Structure of the UEMS: Specialties Sections (and their Divisions)*; 2025. Available from: <https://www.uems.eu/our-structure#Specialties-sections-%26-divisions> [Last accessed on 2025 Aug 18].
59. The Medical Board of Australia. *List of Specialties, Fields of Specialty Practice and Related Specialist Title*; 2025. Available from: <https://www.medicalboard.gov.au/Registration/Types/Specialist-Registration.aspx> [Last accessed on 2025 Aug 18].
60. Kupis R, Michalik B, Polak M, Kulbat M, Domagała A. Specialty choices among new generation of doctors - insights from a Polish survey study. *Sci Rep.* 2024;14(1):27855.
doi: 10.1038/s41598-024-79079-7
61. Borges NJ, Richard GV. Using the delphi method to classify medical specialties. *Career Dev Q.* 2018;66(1):85-90.
doi: 10.1002/cdq.12124
62. Heath JK, Dine CJ, Burke AE, Andolsek KM. Teaching the teachers with milestones: Using the ACGME milestones model for professional development. *J Grad Med Educ.* 2021;13(2 Suppl):124-126.
doi: 10.4300/jgme-d-20-00891.1
63. Porcel JM, Casademont J, Conthe P, Pinilla B, Pujol R, García-Alegría J. Core competencies in Internal Medicine. *Eur J Intern Med.* 2012;23(4):338-341.
doi: 10.1016/j.ejim.2012.03.003
64. Tapia NM, Milewicz A, Whitney SE, Liang MK, Braxton CC. Identifying and eliminating deficiencies in the general surgery resident core competency curriculum. *JAMA Surg.* 2014;149(6):514-518.
doi: 10.1001/jamasurg.2013.4406
65. Relyea-Chew A, Talner LB. A dedicated general competencies curriculum for radiology residents development and implementation. *Acad Radiol.* 2011;18(5):650-654.
doi: 10.1016/j.acra.2010.12.016
66. Popov TA, Ledford D, Lockey R, et al. Maintenance of skills, competencies, and performance in allergy and clinical immunology: Time to lay the foundation for a universal approach. *World Allergy Organ J.* 2012;5(4):45-51.
doi: 10.1097/WOX.0b013e31825546b4
67. Parker NW. Core clinical competencies in anesthesiology: A case based approach. *Anesth Analg.* 2011;112(6):1515.
doi: 10.1213/ANE.0b013e318207b37f
68. Zahid A, Rajan V, Hong J, Young CJ. Surgical competencies required in newly commencing colorectal surgeons: An educational and training spectrum. *Med Sci Educ.* 2020;30(3):1043-1047.
doi: 10.1007/s40670-020-01005-z
69. Kong Kam Wa I, Maule Y, Guerisse F. Key competencies in emergency medicine training in Belgium: A new criteria-based evaluation strategy. *Cureus.* 2024;16(10):e70772.
doi: 10.7759/cureus.70772
70. Massingham LJ, Nuñez S, Bernstein JA, et al. 2022 Association of Professors of Human and Medical Genetics (APHMG) consensus-based update of the core competencies for undergraduate medical education in genetics and genomics. *Genet Med.* 2022;24(10):2167-2179.
doi: 10.1016/j.gim.2022.07.014

71. Giannotta S. Gaining competence in neurological surgery: Is it enough? In: Agarwal N, Reddy V, editors. *Surviving Neurosurgery: Vignettes of Resilience*. Berlin: Springer International Publishing; 2021:263-266.
doi: 10.1007/978-3-030-86917-5_50
72. Jellinger K. Core competencies for neurologists. what clinicians need to know. *Eur J Neurol*. 2004;11:287-287.
doi: 10.1046/j.1468-1331.2003.00760.x
73. Arevalo-Perez J, Paris M, Graham MM, Osborne JR. A perspective of the future of nuclear medicine training and certification. *Semin Nuclear Med*. 2016;46(1):88-96.
doi: 10.1053/j.semnuclmed.2015.10.003
74. Hammond I, Karthigasu K. Training, assessment and competency in gynaecologic surgery. *Best Pract Res Clin Obstet Gynaecol*. 2006;20(1):173-187.
doi: 10.1016/j.bpobgyn.2005.09.006
75. Wentzell DD, Chung H, Hanson C, Gooi P. Competency-based medical education in ophthalmology residency training: A review. *Can J Ophthalmol*. 2020;55(1):12-19.
doi: 10.1016/j.cjco.2019.07.004
76. Kellam J, Archibald D, Barber J, et al. The core competencies for general orthopaedic surgeons. *J Bone Joint Surg*. 2017;99:175-181.
doi: 10.2106/JBJS.16.00761
77. Wiet GJ, Stredney D, Wan D. Training and simulation in otolaryngology. *Otolaryngol Clin North Am*. 2011;44(6):1333-1350, viii-ix.
doi: 10.1016/j.otc.2011.08.009
78. Munson L, Craig LE, Miller MA, et al. Elements of good training in anatomic pathology. *Vet Pathol*. 2010;47(5):995-1002.
doi: 10.1177/0300985810377725
79. Smith BR, Wells A, Alexander CB, et al. Curriculum content and evaluation of resident competency in clinical pathology (laboratory medicine): A proposal. *Human Pathol*. 2006;37(8):934-968.
doi: 10.1016/j.humpath.2006.01.034
80. McIntyre JS. Core competencies for psychiatric practice: What clinicians need to know (a report of the american board of psychiatry and neurology). *Psychiatric Serv*. 2005;56(1):114.
doi: 10.1176/appi.ps.56.1.114
81. Lutz K, Yazdani A, Ross D. From time-based to competency-based standards: Core transitional competencies in plastic surgery. *J Surg Educ*. 2015;72(2):228-34.
doi: 10.1016/j.jsurg.2014.08.013
82. Bethel EC, Marchetti KA, Hecklinski TM, et al. The LEGO™ exercise: An assessment of core competencies in urology residency interviews. *J Surg Educ*. 2021;78(6):2063-2069.
doi: 10.1016/j.jsurg.2021.05.011
83. Popa SL, Ismaiel A, Brata VD, et al. Artificial Intelligence and medical specialties: Support or substitution? *Med Pharm Rep*. 2024;97(4):409-418.
doi: 10.15386/mpr-2696
84. Shang Z, Chauhan V, Devi K, Patil S. Artificial intelligence, the digital surgeon: Unravelling its emerging footprint in healthcare - the narrative review. *J Multidiscip Healthc*. 2024;17:4011-4022.
doi: 10.2147/jmdh.S482757
85. Leivaditis V, Maniatopoulos AA, Lausberg H, et al. Artificial intelligence in thoracic surgery: A review bridging innovation and clinical practice for the next generation of surgical care. *J Clin Med*. 2025;14(8):2729.
doi: 10.3390/jcm14082729
86. Chevalier O, Dubey G, Benkabbou A, Majbar MA, Souadka A. Comprehensive overview of artificial intelligence in surgery: A systematic review and perspectives. *Pflugers Arch*. 2025;477(4):617-626.
doi: 10.1007/s00424-025-03076-6
87. Liu Y, Deng K, Zhang C, Yuan Z, Zhou J, Liu C. The rise of intelligent plastic surgery: A 10-year bibliometric journey through AI applications, challenges, and transformative potential. *Aesthetic Plast Surg*. 2025.
doi: 10.1007/s00266-025-05068-4
88. Lu MY, Chen B, Williamson DFK, et al. A multimodal generative AI copilot for human pathology. *Nature*. 2024;634(8033):466-473.
doi: 10.1038/s41586-024-07618-3
89. Zhang C, Liu S, Zhou X, et al. Examining the role of large language models in orthopedics: Systematic review. *J Med Internet Res*. 2024;26:e59607.
doi: 10.2196/59607
90. Harzing AW. *Publish or Perish*; 2024. Available from: <https://harzing.com/resources/publish-or-perish> [Last accessed on 2024 Oct 01].
91. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *Int J Qual Methods*. 2023;22:16094069231205789.
doi: 10.1177/16094069231205789
92. Ugajin A. Automation in hospitals and health care. In: Nof SY, editor. *Springer Handbook of Automation*. Berlin: Springer International Publishing; 2023:1209-1233.
doi: 10.1007/978-3-030-96729-1_56
93. Munari E, Scarpa A, Cima L, et al. Cutting-edge technology and automation in the pathology laboratory. *Virchows Arch*.

- 2024;484(4):555-566.
doi: 10.1007/s00428-023-03637-z
94. Kalra A, Chakraborty A, Fine B, Reicher J. Machine learning for automation of radiology protocols for quality and efficiency improvement. *J Am Coll Radiol*. 2020;17(9):1149-1158.
doi: 10.1016/j.jacr.2020.03.012
95. Kazzazi F. The automation of doctors and machines: A classification for AI in medicine (ADAM framework). *Future Healthc J*. 2021;8(2):e257-e262.
doi: 10.7861/fhj.2020-0189
96. Rivero-Moreno Y, Rodriguez M, Losada-Muñoz P, et al. Autonomous robotic surgery: Has the future arrived? *Cureus*. 2024;16(1):e52243.
doi: 10.7759/cureus.52243
97. Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technol Forecast Soc Change*. 2017;114:254-280.
doi: 10.1016/j.techfore.2016.08.019
98. Spruit EN, Band GP, Hamming JF, Ridderinkhof KR. Optimal training design for procedural motor skills: A review and application to laparoscopic surgery. *Psychol Res*. 2014;78(6):878-891.
doi: 10.1007/s00426-013-0525-5
99. Ritter FE, Yeh MK, Yan Y, Siu KC, Oleynikov D. Effects of varied surgical simulation training schedules on motor-skill acquisition. *Surg Innov*. 2020;27(1):68-80.
doi: 10.1177/1553350619881591
100. Bould MD, Crabtree NA, Naik VN. Assessment of procedural skills in anaesthesia. *Br J Anaesth*. 2009;103(4):472-483.
doi: 10.1093/bja/aep241
101. Kantak SS, Winstein CJ. Learning-performance distinction and memory processes for motor skills: A focused review and perspective. *Behav Brain Res*. 2012;228(1):219-231.
doi: 10.1016/j.bbr.2011.11.028
102. Levy R, Dubrowski A, Amin H, Bismilla Z. Procedural skills in paediatric residency: Re-evaluating the competencies. *Paediatr Child Health*. 2014;19(4):180-184.
doi: 10.1093/pch/19.4.180
103. Cox CL, Miller BM, Kuhn I, Fritz Z. Diagnostic uncertainty in primary care: What is known about its communication, and what are the associated ethical issues? *Fam Pract*. 2021;38(5):654-668.
doi: 10.1093/fampra/cmab023
104. Scott IA, Doust JA, Keijzers GB, Wallis KA. Coping with uncertainty in clinical practice: A narrative review. *Med J Aust*. 2023;218(9):418-425.
doi: 10.5694/mja2.51925
105. Lockshin MD, Crow MK, Barbhuiya M. When a diagnosis has no name: Uncertainty and opportunity. *ACR Open Rheumatol*. 2022;4(3):197-201.
doi: 10.1002/acr2.11368
106. Kwan JL, Calder LA, Bowman CL, et al. Characteristics and contributing factors of diagnostic error in surgery: Analysis of closed medico-legal cases and complaints in Canada. *Can J Surg*. 2024;67(1):E58-E65.
doi: 10.1503/cjs.003523
107. Howick J, Moscrop A, Mebius A, et al. Effects of empathic and positive communication in healthcare consultations: A systematic review and meta-analysis. *J R Soc Med*. 2018;111(7):240-252.
doi: 10.1177/0141076818769477
108. van Offenbeek MAG, Roemeling OP, Burggraaf A. Exploring professionals' experiences with secure messaging in Dutch outpatient clinics: Emerging differences in use frequencies and types across medical specialties. *BMC Med Inform Decis Mak*. 2025;25(1):245.
doi: 10.1186/s12911-025-03081-w
109. Lakdawala PD. Doctor-patient relationship in psychiatry. *Mens Sana Monogr*. 2015;13(1):82-90.
doi: 10.4103/0973-1229.153308
110. Dugdale DC, Epstein R, Pantilat SZ. Time and the patient-physician relationship. *J Gen Intern Med*. 1999;14 Suppl 1(Suppl 1):S34-S40.
doi: 10.1046/j.1525-1497.1999.00263.x
111. Batko K, Ślęzak A. The use of big data analytics in healthcare. *J Big Data*. 2022;9(1):3.
doi: 10.1186/s40537-021-00553-4
112. Kondylakis H, Kalokyri V, Sfakianakis S, et al. Data infrastructures for AI in medical imaging: A report on the experiences of five EU projects. *Eur Radiol Exp*. 2023;7(1):20.
doi: 10.1186/s41747-023-00336-x
113. Ganslandt T, Neumaier M. Digital networks for laboratory data: Potentials, barriers and current initiatives. *Clin Chem Lab Med*. 2018;57:336-342.
doi: 10.1515/cclm-2018-1131
114. Vincze M, Molnar B, Kozlovsky M. Real-time network video data streaming in digital medicine. *Computers*. 2023;12(11):234.
doi: 10.3390/computers12110234
115. Holman GT, Beasley JW, Karsh BT, Stone JA, Smith PD, Wetterneck TB. The myth of standardized workflow in primary care. *J Am Med Inform Assoc*. 2016;23(1):29-37.
doi: 10.1093/jamia/ocv107
116. Zayas-Cabán T, Okubo TH, Posnack S. Priorities to

- accelerate workflow automation in health care. *J Am Med Inform Assoc.* 2022;30(1):195-201.
doi: 10.1093/jamia/ocac197
117. Beauchemin M, Cohn E, Shelton RC. Implementation of clinical practice guidelines in the health care setting: A concept analysis. *ANS Adv Nurs Sci.* 2019;42(4):307-324.
doi: 10.1097/ans.0000000000000263
118. Pereira VC, Silva SN, Carvalho VKS, Zanghelini F, Barreto JOM. Strategies for the implementation of clinical practice guidelines in public health: An overview of systematic reviews. *Health Res Policy Syst.* 2022;20(1):13.
doi: 10.1186/s12961-022-00815-4
119. Qumseya B, Goddard A, Qumseya A, Estores D, Draganov PV, Forsmark C. Barriers to clinical practice guideline implementation among physicians: A physician survey. *Int J Gen Med.* 2021;14:7591-7598.
doi: 10.2147/ijgm.S333501
120. Thoonsen AC, van Schoten SM, Merten H, *et al.* Stimulating implementation of clinical practice guidelines in hospital care from a central guideline organization perspective: A systematic review. *Health Policy.* 2024;148:105135.
doi: 10.1016/j.healthpol.2024.105135
121. Varkey B. Principles of clinical ethics and their application to practice. *Med Princ Pract.* 2021;30(1):17-28.
doi: 10.1159/000509119
122. Raveesh BN, Nayak RB, Kumbar SF. Preventing medico-legal issues in clinical practice. *Ann Indian Acad Neurol.* 2016;19(Suppl 1):S15-S20.
doi: 10.4103/0972-2327.192886
123. Garcia JR, Allende F, Kogan M, Chahla J. When things go wrong: A guide to the medical, ethical, and legal dimensions of surgical complications. *J Am Acad Orthop Surg Glob Res Rev.* 2024;8(12):e24.00004.
doi: 10.5435/JAAOSGlobal-D-24-00004
124. Owen G, Kanaan R. The legal and ethical framework for psychiatry. *Medicine.* 2008;36:391-392.
doi: 10.1016/j.mpm.2008.05.010
125. Waller L. Ethics, law and paediatric medicine. *J Paediatr Child Health.* 2011;47(9):620-623.
doi: 10.1111/j.1440-1754.2011.02167.x
126. Coulter C, McKay F, Hallowell N, *et al.* Understanding the ethical and legal considerations of Digital Pathology. *J Pathol Clin Res.* 2022;8(2):101-115.
doi: 10.1002/cjp.2.251
127. Harvey-Lloyd JM. Operating within the legal and ethical framework to gain co-operation when imaging paediatric patients. *Radiography.* 2013;19(4):285-289.
doi: 10.1016/j.radi.2013.07.010
128. Sallam M, Barakat M, Sallam M. A preliminary checklist (METRICS) to standardize the design and reporting of studies on generative artificial intelligence-based models in health care education and practice: Development study involving a literature review. *Interact J Med Res.* 2024;13:e54704.
doi: 10.2196/54704
129. Najjar R. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics (Basel).* 2023;13(17):2760.
doi: 10.3390/diagnostics13172760
130. O'Brien AJ, Vrazas JJ, Clements W. Current applications of algorithmic artificial intelligence in interventional radiology: A review of the literature. *J Med Imaging Radiat Oncol.* 2024;68(2):194-207.
doi: 10.1111/1754-9485.13609
131. Battaglia E, Boehm J, Zheng Y, Jamieson AR, Gahan J, Majewicz Fey A. Rethinking autonomous surgery: Focusing on enhancement over autonomy. *Eur Urol Focus.* 2021;7(4):696-705.
doi: 10.1016/j.euf.2021.06.009
132. Schmidgall S, Opfermann JD, Kim JW, Krieger A. Will your next surgeon be a robot? Autonomy and AI in robotic surgery. *Sci Robot.* 2025;10(104):eadt0187.
doi: 10.1126/scirobotics.adt0187
133. Eloy C, Fraggetta F, van Diest PJ, *et al.* Digital transformation of pathology - the European Society of Pathology expert opinion paper. *Virch Archiv.* 2025.
doi: 10.1007/s00428-025-04090-w
134. Khavandi S, Zaghloul F, Higham A, Lim E, de Pennington N, Celi LA. Investigating the impact of automation on the health care workforce through autonomous telemedicine in the cataract pathway: Protocol for a multicenter study. *JMIR Res Protoc.* 2023;12:e49374.
doi: 10.2196/49374
135. Hirani R, Noruzi K, Khuram H, *et al.* Artificial intelligence and healthcare: A journey through history, present innovations, and future possibilities. *Life (Basel).* 2024;14(5):557.
doi: 10.3390/life14050557
136. Schuitmaker L, Drogts J, Benders M, Jongsma K. Physicians' required competencies in AI-assisted clinical settings: A systematic review. *Br Med Bull.* 2025;153(1):ldae025.
doi: 10.1093/bmb/ldae025
137. Schlemmer HP. Navigating the AI revolution: Will radiology sink or soar? *Jpn J Radiol.* 2025;43:1628-1633.
doi: 10.1007/s11604-025-01810-9

138. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500-510.
doi: 10.1038/s41568-018-0016-5
139. Berbís MA, McClintock DS, Bychkov A, *et al.* Computational pathology in 2030: A Delphi study forecasting the role of AI in pathology within the next decade. *EBioMedicine*. 2023;88:104427.
doi: 10.1016/j.ebiom.2022.104427
140. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16(11):703-715.
doi: 10.1038/s41571-019-0252-y
141. Pallua JD, Brunner A, Zelger B, Schirmer M, Haybaeck J. The future of pathology is digital. *Pathol Res Pract*. 2020;216:153040.
doi: 10.1016/j.prp.2020.153040
142. Atabansi CC, Nie J, Liu H, Song Q, Yan L, Zhou X. A survey of Transformer applications for histopathological image analysis: New developments and future directions. *Biomed Eng Online*. 2023;22(1):96.
doi: 10.1186/s12938-023-01157-0
143. Alsulimani A, Akhter N, Jameela F, *et al.* The impact of artificial intelligence on microbial diagnosis. *Microorganisms*. 2024;12(6):1051.
doi: 10.3390/microorganisms12061051
144. Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. *Front Public Health*. 2023;11:1273253.
doi: 10.3389/fpubh.2023.1273253
145. Vandenberg O, Durand G, Hallin M, *et al.* Consolidation of clinical microbiology laboratories and introduction of transformative technologies. *Clin Microbiol Rev*. 2020;33(2):e00057-19.
doi: 10.1128/cmr.00057-19
146. Sahu P, Kumar S, Kumar N, Prakash V. Total laboratory automation in clinical microbiology: A note on needs, challenges, and applications in a pandemic scenario. *Appl Microbiol Theory Technol*. 2024;5:64-71.
doi: 10.37256/amtt.5120243753
147. Spies NC, Rangel A, English P, Morrison M, O'Fallon B, Ng DP. Machine learning methods in clinical flow cytometry. *Cancers (Basel)*. 2025;17(3):483.
doi: 10.3390/cancers17030483
148. de Almeida JG, Gudgin E, Besser M, *et al.* Computational analysis of peripheral blood smears detects disease-associated cytomorphologies. *Nat Commun*. 2023;14(1):4378.
doi: 10.1038/s41467-023-39676-y
149. Lin Y, Mensah IK, Doering M, Shean RC, Spies NC. Machine learning-based error detection in the clinical laboratory: A critical review. *Crit Rev Clin Lab Sci*. 2025:1-13.
doi: 10.1080/10408363.2025.2512468
150. Akinrinmade AO, Adebile TM, Ezuma-Ebong C, *et al.* Artificial intelligence in healthcare: Perception and reality. *Cureus*. 2023;15(9):e45594.
doi: 10.7759/cureus.45594
151. Blease C, Locher C, Leon-Carlyle M, Doraiswamy M. Artificial intelligence and the future of psychiatry: Qualitative findings from a global physician survey. *Digit Health*. 2020;6:2055207620968355.
doi: 10.1177/2055207620968355
152. Alarcón RD. Culture, cultural factors and psychiatric diagnosis: Review and projections. *World Psychiatry*. 2009;8(3):131-139.
doi: 10.1002/j.2051-5545.2009.tb00233.x
153. Ni Y, Jia F. A scoping review of AI-driven digital interventions in mental health care: Mapping applications across screening, support, monitoring, prevention, and clinical education. *Healthcare (Basel)*. 2025;13(10):1205.
doi: 10.3390/healthcare13101205
154. Ficerai-Garland G, Groves P, Puccio EA, *et al.* Caregiver and pediatric clinician perspectives on artificial intelligence for language services. *Acad Pediatr*. 2025:102887.
doi: 10.1016/j.acap.2025.102887
155. Kurath-Koller S. Artificial intelligence in pediatrics: Promise, peril, and the path ahead. *Front Pediatr*. 2025;13:1631521.
doi: 10.3389/fped.2025.1631521
156. Katsakiori PF, Kagadis GC, Mulita F, Marangos M. Implementing artificial intelligence in family medicine: Challenges and limitations. *Cureus*. 2024;16(12):e75518.
doi: 10.7759/cureus.75518
157. Hanna K, Chartash D, Liaw W, *et al.* Family medicine must prepare for artificial intelligence. *J Am Board Fam Med*. 2024;37(4):520-524.
doi: 10.3122/jabfm.2023.230360R1
158. Ventres WB, Stone LA, Joslin TA, *et al.* Storylines of family medicine III: Core principles-primary care, systems and family. *Fam Med Community Health*. 2024;12(Suppl 3):e002790.
doi: 10.1136/fmch-2024-002790
159. Yu S, El Atrache R, Tang J, *et al.* Artificial intelligence-enhanced epileptic seizure detection by wearables. *Epilepsia*. 2023;64(12):3213-3226.
doi: 10.1111/epi.17774

160. Khalifa M, Albadawy M. AI in diagnostic imaging: Revolutionising accuracy and efficiency. *Computer Methods and Programs in Biomedicine Update*. 2024;5:100146.
doi: 10.1016/j.cmpbup.2024.100146
161. Kalani M, Anjankar A. Revolutionizing neurology: The role of artificial intelligence in advancing diagnosis and treatment. *Cureus*. 2024;16(6):e61706.
doi: 10.7759/cureus.61706
162. Shah U, Alzubaidi M, Mohsen F, Abd-alrazaq A, Alam T, Househ M. The role of artificial intelligence in decoding speech from EEG signals: A scoping review. *Sensors (Basel)*. 2022;22(18):6975.
doi: 10.3390/s22186975
163. Olawade DB, Wada OJ, David-Olawade AC, Kunonga E, Abaire O, Ling J. Using artificial intelligence to improve public health: A narrative review. *Front Public Health*. 2023;11:1196397.
doi: 10.3389/fpubh.2023.1196397
164. Sallam M. The fragility of public health in the face of controversial leadership. *BMJ Glob Health*. 2025;10(7):e018536.
doi: 10.1136/bmjgh-2024-018536
165. Sumner J, Lim H, Chong W, Bunde A, Mukhopadhyay A, Kayambu G. Artificial intelligence in physical rehabilitation: A systematic review. *Artif Intell Med*. 2023;146:102693.
doi: 10.1016/j.artmed.2023.102693
166. MohammadNamdar M, Lowery Wilson M, Murtonen KP, Aartolahti E, Oduor M, Korniloff K. How AI-based digital rehabilitation improves end-user adherence: Rapid review. *JMIR Rehabil Assist Technol*. 2025;12:e69763.
doi: 10.2196/69763
167. Amin A, Cardoso SA, Suyambu J, et al. Future of artificial intelligence in surgery: A narrative review. *Cureus*. 2024;16(1):e51631.
doi: 10.7759/cureus.51631
168. Lee R, Ziehm A, Ullrich L, Stawicki S. *Artificial Intelligence in Surgery, Surgical Subspecialties, and Related Disciplines*. London: Intechopen; 2023.
169. Shah M, Naik N, Somani BK, Hameed BMZ. Artificial intelligence (AI) in urology-Current use and future directions: An iTRUE study. *Turk J Urol*. 2020;46(Suppl. 1):S27-S39.
doi: 10.5152/tud.2020.20117
170. Ghneim MH, O'Connor JV, Scalea TM. Damage control thoracic surgery: What you need to know. *J Trauma Acute Care Surg*. 2025;98(1):11-19.
doi: 10.1097/ta.0000000000004458
171. Zhou X, Chen Y, Miao G, Guo Y, Zhang Q, Bi J. Computer-aided robotics for applications in fracture reduction surgery: Advances, challenges, and opportunities. *iScience*. 2025;28(1):111509.
doi: 10.1016/j.isci.2024.111509
172. Choi J. Cosmetic surgery: Is it science or art? *Arch Plast Surg*. 2015;42:672-674.
doi: 10.5999/aps.2015.42.5.672
173. Patel DJ, Chaudhari K, Acharya N, Shrivastava D, Muneeba S. Artificial intelligence in obstetrics and gynecology: Transforming care and outcomes. *Cureus*. 2024;16(7):e64725.
doi: 10.7759/cureus.64725
174. Tangsrivimol JA, Schonfeld E, Zhang M, et al. Artificial intelligence in neurosurgery: A state-of-the-art review from past to future. *Diagnostics (Basel)*. 2023;13(14):2429.
doi: 10.3390/diagnostics13142429
175. Berg M, Dahlin LB, Kjellman M. Overview of surgical training and assessment of surgical skills: A narrative review. *Front Surg*. 2025;12:1605495.
doi: 10.3389/fsurg.2025.1605495
176. Pakkasjärvi N, Anttila H, Pyhältö K. What are the learning objectives in surgical training – a systematic literature review of the surgical competence framework. *BMC Med Educ*. 2024;24(1):119.
doi: 10.1186/s12909-024-05068-z
177. Levkovich I. Is artificial intelligence the next copilot for primary care in diagnosing and recommending treatments for depression? *Med Sci (Basel)*. 2025;13(1):8.
doi: 10.3390/medsci13010008
178. Behara K, Bhero E, Agee JT. AI in dermatology: A comprehensive review into skin cancer detection. *PeerJ Comput Sci*. 2024;10:e2530.
doi: 10.7717/peerj-cs.2530
179. Lim JI, Regillo CD, Sadda SR, et al. Artificial intelligence detection of diabetic retinopathy: Subgroup comparison of the eyeart system with ophthalmologists' dilated examinations. *Ophthalmol Sci*. 2023;3(1):100228.
doi: 10.1016/j.xops.2022.100228
180. Li Z, Wang L, Wu X, et al. Artificial intelligence in ophthalmology: The path to the real-world clinic. *Cell Rep Med*. 2023;4(7):101095.
doi: 10.1016/j.xcrm.2023.101095
181. Amiot F, Potier B. Artificial intelligence (AI) and emergency medicine: Balancing opportunities and challenges. *JMIR Med Inform*. 2025;13:e70903.
doi: 10.2196/70903
182. Bellini V, Rafano Carnà E, Russo M, et al. Artificial intelligence and anesthesia: A narrative review. *Ann Transl Med*. 2022;10(9):528.

- doi: 10.21037/atm-21-7031
183. Xiao N, Huang X, Wu Y, *et al.* Opportunities and challenges with artificial intelligence in allergy and immunology: A bibliometric study. *Front Med (Lausanne)*. 2025;12:1523902.
doi: 10.3389/fmed.2025.1523902
184. MacMath D, Chen M, Khoury P. Artificial intelligence: Exploring the future of innovation in allergy immunology. *Curr Allergy Asthma Rep*. 2023;23(6):351-362.
doi: 10.1007/s11882-023-01084-z
185. Nair A, Ong W, Lee A, *et al.* Enhancing radiologist productivity with artificial intelligence in magnetic resonance imaging (MRI): A narrative review. *Diagnostics (Basel)*. 2025;15(9):1146.
doi: 10.3390/diagnostics15091146
186. Korfiatis P, Kline TL, Meyer HM, *et al.* Implementing artificial intelligence algorithms in the radiology workflow: Challenges and considerations. *Mayo Clin Proc Digit Health*. 2025;3(1):100188.
doi: 10.1016/j.mcpdig.2024.100188
187. Hirata K, Sugimori H, Fujima N, Toyonaga T, Kudo K. Artificial intelligence for nuclear medicine in oncology. *Ann Nuclear Med*. 2022;36:123-132.
doi: 10.1007/s12149-021-01693-6
188. Odonkor B, Kaggwa S, Uwaoma P, Hassan A, Farayola O. The impact of AI on accounting practices: A review: Exploring how artificial intelligence is transforming traditional accounting methods and financial reporting. *World J Adv Res Rev*. 2024;21:172-188.
doi: 10.30574/wjarr.2024.21.1.2721
189. Khakpaki A. Advancements in artificial intelligence transforming medical education: A comprehensive overview. *Med Educ Online*. 2025;30(1):2542807.
doi: 10.1080/10872981.2025.2542807
190. Pauley K, Flin R, Yule S, Youngson G. Surgeons' intraoperative decision making and risk management. *Am J Surg*. 2011;202(4):375-381.
doi: 10.1016/j.amjsurg.2010.11.009
191. Cristancho SM, Apramian T, Vanstone M, *et al.* Thinking like an expert: Surgical decision making as a cyclical process of being aware. *Am J Surg*. 2016;211(1):64-69.
doi: 10.1016/j.amjsurg.2015.03.010
192. Ardito RB, Rabellino D. Therapeutic alliance and outcome of psychotherapy: Historical excursus, measurements, and prospects for research. *Front Psychol*. 2011;2:270.
doi: 10.3389/fpsyg.2011.00270
193. Aramrat C, Choksomngam Y, Jiraporncharoen W, *et al.* Advancing multimorbidity management in primary care: A narrative review. *Prim Health Care Res Dev*. 2022;23:e36.
doi: 10.1017/s1463423622000238
194. Tekin M, Yurdal MO, Toraman Ç, Korkmaz G, Uysal İ. Is AI the future of evaluation in medical education?? AI vs. human evaluation in objective structured clinical examination. *BMC Med Educ*. 2025;25(1):641.
doi: 10.1186/s12909-025-07241-4
195. Cestonaro C, Delicati A, Marcante B, Caenazzo L, Tozzo P. Defining medical liability when artificial intelligence is applied on diagnostic algorithms: A systematic review. *Front Med (Lausanne)*. 2023;10:1305756.
doi: 10.3389/fmed.2023.1305756

Appendix 1

GenAI prompt used

You are to score the risk of displacement of each medical specialty by superintelligent AI.

Follow these instructions exactly, without deviation. Do not assume additional context beyond what is given.

1. Task

For each medical specialty provided, assign seven dimension scores on a 0–5 scale according to the definitions below. Then, apply reverse scoring to specific dimensions, sum the adjusted values, and categorize the total risk.

2. Scoring Scale

- 0. = Not present or negligible in the specialty. 1 = Very low.
- 2. = Low.
- 3. = Moderate.
- 4. = High.
- 5. = Very high.

3. Dimension Definitions

#A. Risk-Increasing Dimensions (higher score increases displacement risk; no reverse scoring)

- 1. Task Automatability – How easily the specialty’s core tasks could be automated by AI within the next 10–15 years.
 - 0. = No tasks automatable.
 - 5. = Nearly all tasks can be automated.
- 2. Data-Centricity – Degree to which the specialty relies on structured, digitized, and AI-compatible data.
 - 0. = Almost no structured data.
 - 5. = Highly dependent on structured datasets.
- 3. Standardizability – Extent to which workflows are uniform, protocol-driven, and reproducible across contexts.
 - 0. = Completely individualized.
 - 5. = Fully standardized across cases.

B. Risk-Mitigation Dimensions (higher raw score lowers displacement risk; reverse scoring required)

- 4. Procedural Complexity – Level of intricate, high-skill manual execution required. 0 = No manual skill.

- 5. = Highly intricate multi-step manual procedures.

- 5. Diagnostic Ambiguity – Amount of uncertainty and complexity in reaching a diagnosis requiring nuanced human judgment.

- 0. = Fully objective, binary diagnoses.

- 5. = Highly ambiguous, multifactorial diagnoses.

- 6. Patient Interaction Intensity – Frequency and centrality of direct patient communication, counseling, and trust-building.

- 0. = No patient interaction.

- 5. = Constant and critical patient engagement.

- 7. Ethical/Legal Complexity – Extent to which the specialty involves high-stakes ethical or legal decision-making.

- 0. = Minimal ethical/legal issues.

- 5. = Extremely high-stakes ethical/legal contexts.

4. Reverse Scoring

After assigning raw scores for all seven dimensions:

For Procedural Complexity, Diagnostic Ambiguity, Patient Interaction Intensity, and Ethical/Legal Complexity, calculate:

Adjusted Score = 5 – Raw Score

Replace the raw score with the adjusted score before summing totals.

5. Total Risk Score

- 1. Keep Task Automatability, Data-Centricity, and Standardizability scores unchanged.
- 2. Use adjusted scores for the other four dimensions.
- 3. Sum all seven adjusted scores. The range is 0–35.

6. Output Format

For each specialty, output in this exact table format:

\| Specialty | Task Automatability | Procedural Complexity (raw) | Diagnostic Ambiguity (raw) | Patient Interaction Intensity (raw) | Data-Centricity | Standardizability | Ethical/Legal Complexity (raw) | Procedural Complexity (adjusted) | Diagnostic Ambiguity (adjusted) | Patient Interaction Intensity (adjusted) | Ethical/Legal Complexity (adjusted) | Total Score | Risk Category |