Review Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2026.65.010141

Training Programs for Medical Professionals to Use with Language Models

ON Andreeva[1], EV Malkova[2], AA Ryazanova[3], AA Shcherbakova[4] and A Yu Shcherbakov[5]*

[1]*Ph.D of Medical Sciences, doctor of telemedicine consultations, Head of the Gynecological Office Doctor 2000 LLC (Partner of the network of medical centers "Med Swiss"), France*

[2]*Head of the Scientific and Educational Center of Social Analytics of the Russian State Social University, Russia*

[3]*Leading Specialist of the Scientific and Educational Center of Social Analytics of the Russian State Social University, Russia*

[4]*Independent researcher, Marseille, France*

[5]*Grand Ph.D., Professor of Technical Sciences, Head of the Department of Cognitive-Analytical and Neuro-Applied Technologies of the Russian State Social University, Leading Researcher of the State University of Management, Russia*

***Corresponding author:** A Yu Shcherbakov, Grand Ph.D., Professor of Technical Sciences, Head of the Department of Cognitive-Analytical and Neuro-Applied Technologies of the Russian State Social University, Leading Researcher of the State University of Management, Russia

ARTICLE INFO

Received: March 02, 2026

Published: March 20, 2026

Citation: ON Andreeva, EV Malkova, AA Ryazanova, AA Shcherbakova and A Yu Shcherbakov. Training Programs for Medical Professionals to Use with Language Models. Biomed J Sci & Tech Res 65(1)-2026. BJSTR. MS.ID.010141.

ABSTRACT

Large language models (LLMs) such as ChatGPT and related generative systems are rapidly entering clinical workflows, medical education, and biomedical research, creating an urgent need to train medical professionals in their safe and effective use [1-4]. This article proposes a conceptual framework for training programs aimed at clinicians, residents, nurses, and other health professionals, addressing core competencies, curriculum structure, pedagogical strategies, and assessment approaches. Drawing on emerging literature in medical education and clinician-facing guidance, we argue that training must go beyond basic tool familiarization to cover critical appraisal of model outputs, data protection, regulatory compliance, and human–AI collaboration skills [1-3,5,6]. We outline a multi-module program that integrates foundational AI literacy, hands-on clinical scenarios, safety and ethics training, and institution-specific governance, and we discuss implementation challenges and future directions.

Abbreviations: LLMs: Large language Models; PBL: Problem-Based Learning; CBL: Case-Based Learning; OSCEs: Objective Structured Clinical Examinations

Introduction

Generative LLMs are poised to influence nearly all aspects of medicine, from clinical documentation and decision support to patient communication and continuous professional development [1,2,4,7]. Early work in medical education shows that integrating LLMs into courses can enhance problem-based learning, support clinical reasoning, and provide personalized learning resources, while also introducing new forms of cognitive offloading and automation bias [1-3,8]. At the same time, professional bodies and regulators emphasize that AI tools must be used under strong clinical oversight, with attention to privacy, personal data treatment transparency and security, explainability, and accountability [5,6,9]. Despite this rapid diffusion, many clinicians report limited understanding of how LLMs work, their limitations, and the governance structures surrounding their use [4,10,11]. Existing offerings, such as short online courses on foundations of LLMs in biomedical research, demonstrate high demand but usually provide only an introductory overview [12]. There is a pressing need for structured training programs that develop durable competencies for safe and critical use of LLMs in everyday clinical practice. This article addresses the following questions:

1. What competencies should medical professionals acquire to work effectively with LLMs?
2. How can training programs be structured to support these competencies in different professional groups?
3. Which pedagogical strategies and assessment methods are most suitable for teaching LLM-related skills?

Background: LLMs in Medicine and Medical Education

LLMs, as a class of large-scale foundation models, are trained on massive text corpora and can generate coherent, contextually appropriate responses to prompts, enabling applications in summarization, question answering, translation, and interactive tutoring [1,2]. In medicine, they are being explored for drafting discharge summaries, simplifying patient information, supporting diagnostic reasoning, and assisting in literature review and evidence synthesis [1,2,4,7]. In medical education, several authors have argued that LLMs can transform curriculum delivery and assessment by enabling dynamic case simulation, automated feedback, and adaptive learning pathways [1-3,8]. For instance, studies integrating ChatGPT into problem-based and case-based learning reported high student acceptance and perceived value when the tool was embedded into course design, rather than offered as an optional add-on [2,3]. At the same time, systematic reviews of the early literature underline concerns about hallucinated content, lack of source transparency, and risk of plagiarism and over-reliance on AI-generated text [1,8]. Guidance documents for general practice and for AI conversational agents in health care converge on several core principles: AI outputs must be double-checked, patient data should not be entered into general-purpose chatbots, clinicians remain responsible for decisions, and use of AI should be transparent to patients with appropriate consent [5,6,9]. Training programs for medical professionals must therefore integrate not only technical skills but also professional norms and regulatory requirements.

Core Competencies for Medical Professionals Working with LLMs

Drawing on narrative reviews of skill shifts in an LLM-enabled health system and clinician-oriented guides, we outline a competency model spanning technical, cognitive, ethical, and organizational dimensions [1,3,4,10,11].

Foundational AI and LLM Literacy

Medical professionals require a basic understanding of how LLMs are trained, what types of data they use, and what "probabilistic text generation" implies for reliability and uncertainty [1,4,10]. This includes familiarity with concepts such as training data bias, fine-tuning, prompt engineering, and the distinction between pattern recognition and domain knowledge [1,4,10,13].

Critical Appraisal and Clinical Reasoning with LLMs

Several authors emphasize that LLMs should be seen as tools that support, rather than replace, clinical reasoning [2-4,7], therefore required clinicians' skills are:

- Formulate clear, structured prompts aligned with clinical questions [1,10,12].
- Interpret outputs critically, identifying hallucinations, inconsistencies, and lack of justification [1,4,10].
- Integrate LLM suggestions with evidence-based guidelines and patient-specific data, maintaining independent judgment [2,5-7].

Data Protection, Privacy, and Regulatory Compliance

Professional guidance stresses strict limits on sharing identifiable patient data with general-purpose AI tools and highlights obligations related to privacy, consent, and documentation [5,6,9]. Clinicians must understand:

- What kinds of data may be entered into institution-approved LLM tools and under which conditions
- How legal frameworks (e.g., health data protection laws and AI-specific regulations) affect LLM use in documentation and decision support [5,6,9].
- How to communicate AI use to patients and document AI-assisted components of care [5,6].

Ethics, Professionalism, and Equity

Ethical guidance for AI conversational agents highlights transparency, informed consent, mitigation of bias, and clear escalation pathways for high-risk situations [5,6,9]. Potential training should help clinicians to recognize:

- Risks of bias and unequal performance across different populations and languages.
- Automation bias and over-trust in apparently confident outputs.
- The importance of explaining the role and limitations of AI tools to patients [5,6,9].

Human–AI Collaboration and Workflow Integration

Clinicians must develop practical competencies in incorporating LLMs into real workflows, including documentation, information retrieval, patient education materials, and communication with colleagues [3,7,10,11]. Evident benefits of machine and deep learning use for early detection of various diseases, in particularly, cancers using image analysis and pattern recognition are well reported in literature [14]. Following this application, the analysis of the integrity of patient's complaints and other narrative data previously shared with AI tools could be promising in diagnosing the severe and potentially lethal diseases at early stages, particularly in contexts where continuous healthcare access could not be provided (i.e. remote and rural areas and communities). Strict interdiction to use this type of personal data in general-purpose chatbots should still be followed but protected tools integrated with telemedicine platforms and digital documentation systems in healthcare institutions could be developed to collect and analyse patients' narrative data and clinical records in order to make the practitioners pay a specific attention to the warning cases. Yet, the full potential of incorporating LLMs into real workflows

requires deep understanding which tasks are appropriate for LLM support and which require traditional tools or specialist consultation along with the knowledge of specific data protection and AI-regulation legal frameworks.

Proposed Structure of Training Programs

Training programs for medical professionals should be modular and adaptable to different professional roles, levels of prior expertise, and institutional contexts [3,4,10-12]. Table 1 summarizes a possible module structure.

Table 1: Modular structure for LLM training programs for medical professionals.

Module	Focus	Key content
1. Foundations of LLMs	Basic literacy	What LLMs are, capabilities and limitations, examples in medicine [1,2,4].
2. Clinical Use Cases	Applications	Documentation support, patient communication, differential diagnosis assistance, literature research [1-3,7].
3. Prompting and Evaluation	Interaction skills	Prompt engineering, iterative querying, crosschecking outputs, recognizing hallucinations [1,3,10,12].
4. Safety, Privacy, Regulation	Governance	Data protection, institutional policies, legal frameworks, accountability [5,6,9].
5. Ethics and Professionalism	Values	Transparency with patients, bias awareness, automation bias, escalation strategies [5,6,9].
6. Clinical Integration	Workflow	Embedding LLMs into documentation, handovers, telemedicine, multidisciplinary teamwork [3,7,10,11].
7. Continuous Learning	Updating	Monitoring tool updates, participating in institutional governance, reflecting on practice [4,8,10,11].

Introductory and Advanced Tracks

Programs may distinguish an introductory track for general users and an advanced track for "clinical AI champions" who support local implementation and policy [4,10-12]. Introductory courses can follow the model of short self-paced online modules with a final assessment and certificate, as seen in existing offerings on LLM foundations in medical research [12]. Advanced tracks could cover evaluation of LLM tools, participation in procurement and risk assessment, and collaboration with data science teams [6,4,10,11].

Tailoring to Professional Roles

Acquired skills need vary between medical students, residents, practicing clinicians, nurses, and allied health professionals, as well as between research-oriented and primarily clinical roles [1,3,8,10]. For example, students may focus on using LLMs for learning and clinical reasoning exercises, whereas attending physicians may emphasize documentation efficiency and risk management in high-stakes decision support [1-3,7].

Pedagogical Strategies

Evidence from initial implementations suggests that LLM training is most effective when embedded into existing curricula and clinical education structures rather than taught in isolation [1-3,8]. Several pedagogical strategies appear particularly promising.

Problem-Based and Case-Based Learning with LLMs

Studies integrating ChatGPT into problem-based learning (PBL) and case-based learning (CBL) report that students value the opportunity to interact with AI while teachers guide critical evaluation of outputs [2,3]. Cases can be designed to teach students to:

- Use an LLM to generate differential diagnoses, then compare them with guideline-based reasoning.
- Ask the model to draft patient explanations and then refine them for clarity and appropriateness in diverse cultural and social contexts.
- Critique AI-generated content for accuracy, completeness, and transparency of reasoning [2,3,7].

Experiential Learning and "Stress Testing"

Narrative reviews highlight the importance of experiencing both the strengths and failure modes of LLMs to build robust mental models [4,10,11]. Training can include exercises where participants deliberately probe the model with ambiguous prompts, rare conditions, or incomplete information to observe hallucinations, over-confidence, and bias, followed by structured reflection.

Explicit Teaching of Safety and Governance

Because safety and data protection remain central concerns in LLM use in medicine, dedicated sessions should focus on institutional policies, external guidelines, and practical "dos and don'ts" of AI use in clinical practice [5,6,9]. For example, general practice guidance emphasizes that sensitive patient data should never be entered into public conversational AI tools and that clinicians must double-check AI outputs [5].

Interprofessional and Team-Based Training

LLMs will affect not only individual clinicians but also team workflows and patient pathways [3,7,10,11]. Interprofessional and interdisciplinary workshops involving physicians, nurses, pharmacists, and administrative staff can explore shared scenarios such as AI-assisted documentation or triage chatbots, clarifying responsibilities and escalation procedures.

Assessment of Competencies

Robust assessment strategies are required to ensure that participants acquire not only declarative knowledge but also practical and ethical competencies [1,3,8,10].

Knowledge Tests and Self-Assessment

Short quizzes and summative exams, similar to those used in existing online courses, can assess understanding of AI concepts, safety principles, and institutional policies [1,10,12]. Self-assessment questionnaires may capture confidence and perceived preparedness to use LLMs in practice [3,10].

Performance-Based Assessment

Objective structured clinical examinations (OSCEs) and workplace-based assessments can be adapted to include AI components, such as evaluating a student's ability to use an LLM to prepare for a patient encounter while maintaining privacy and critically appraising outputs [1,3,8]. Performance can be judged on prompt formulation, detection of errors in AI suggestions, and integration with clinical reasoning.

Reflective Portfolios

Given the rapidly evolving nature of LLM technology, reflective portfolios documenting actual use cases, decisions to refrain from using AI, and personal learning goals may provide a flexible tool for ongoing assessment [4,8,10]. Such portfolios can also support institutional learning about how LLMs are being adopted in practice.

Implementation Challenges

Developing and sustaining effective training programs involves several challenges that institutions should be able to address, continuously and with agility.

Rapid Technological Change

LLM capabilities, interfaces, and deployment models are evolving quickly, making static training materials obsolete [8,4,10,11]. Programs should be provided with mechanisms for continuous updating, possibly through micro-learning units, periodic updates, and close collaboration between educators, informatics teams, and AI governance committees [10,11].

Heterogeneity of Tools and Governance

Institutions' approaches vary widely in regards of which LLM tools are approved, how they are integrated into electronic health records, and what governance structures exist [5,6,9,11]. Training must be highly contextualized to local policies and infrastructure, which requires alignment between education, IT, compliance, and clinical leadership.

Equity and Access

There is a risk that only well-resourced institutions or early adopters will offer high-quality LLM training, potentially widening disparities in digital competencies and patient care [4,8,10]. Open educational resources and cross-institutional collaborations may help in reducing this gap and the challenge of equal access and proficiency of AI tools across institutions and practitioners of different social and geographic backgrounds should be address at the earliest stages.

Future Directions

Future research should evaluate the impact of LLM training programs on clinical practice and patient outcomes, not merely on self-reported comfort with AI tools [3,8,10,11]. Longitudinal studies could examine how training influences patterns of AI use, error detection, and incident reporting, as well as its effects on burnout and documentation burden [10,11]. Another priority is to develop competency frameworks and accreditation standards that can guide curriculum design across institutions and countries [4,8]. Finally, as specialized medical LLMs and multimodal models become more widely available, training will need to incorporate model comparison, validation of domain-specific tools, and a deeper engagement with issues of transparency, interpretability, and shared decision-making with patients [6,4,10,11].

Conclusion

Training programs for medical professionals to work with language models must extend beyond introductory AI literacy to cultivate critical appraisal skills, ethical awareness, and practical competencies for integrating LLMs into clinical workflows. Emerging evidence from medical education and professional guidance suggests that embedded, case-based, and governance-aligned training can help clinicians harness the benefits of LLMs while maintaining patient safety and professional accountability [1-6,10]. Developing robust curricula, assessments, and institutional support will be essential for evaluating the potential of LLMs in healthcare.

References

1. Alaa Abd alrazaq, Rawan AlSaad, Dari Alhuwail, Arfan Ahmed, Padraig Mark Healy, et al. (2023) Large language models in medical education: Opportunities, challenges, and future directions. JMIR Medical Education 9: e48291.

2. Jiang Y, Chen H, Zhang X, Yuqi Yang, Kai Kang, et al (2024) Embracing ChatGPT for medical education: Exploring its impact on doctors and medical students. Cureus 16(4): e58123.
3. Anita V Thomae, Claudia M Witt, Jürgen Barth (2024) Integration of ChatGPT into a course for medical students: Design, implementation, and evaluation. Medical Education Online 10: e50545.
4. Heston TF, Khun C, Bailey AM (2025) The clinicians' guide to large language models: A general framework for safe, effective, and evidence-based implementation. Interactive Journal of Medical Research 14: e59823.
5. (2025) AGPAL, Artificial intelligence in general practice: Industry guidance for safe and smart use.
6. (2025) USC Institute for Creative Technologies, Ensuring safety, ethics, and efficacy: Best practices for AI conversational agents in healthcare.
7. (2025) Medizin Online, Dr. ChatGPT: Large language models in everyday clinical practice.
8. Kumar V, Shah S, Patel R (2024) Thematic foci in large language models research for medical education. BMC Medical Education 24: 1248.
9. (2026) World-first safety guide for public use of AI health chatbots. University of Birmingham.
10. Chen L, Davis M, Evans R (2024) Understanding and training for the impact of large language models in healthcare. JAMA Health Forum 5(10): e244567.
11. Patel S, Johnson K, Williams A (2025) Implementing large language models in health care: Clinician perspectives and training needs. Journal of Medical Internet Research 27: e71916.
12. (2025) Mayo Clinic Executive Education. Foundations of large language models in medical research.
13. (2024) Training large language models: Techniques and best practices. Nitor Infotech.
14. Shaista Suhail, Devesh K Yadav, Fahad M Samadi Alfisha Fatima (2026) AI -A Novel Tool for Oral Cancer Diagnosis and Prevention: Current Applications and Future Perspectives. Am J Biomed Sci & Res 30(2): 1138-1143.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2026.65.010141

A Yu Shcherbakov. Biomed J Sci & Tech Res

Submission Link: https://biomedres.us/submit-manuscript.php