NEJM AI 2025;2(5)
DOI: 10.1056/AIp2401059

PERSPECTIVE

Regulation of AI: Learnings from Medical Education

Kerstin N. Vokinger, M.D., J.D., Ph.D., L.L.M.,[1] Derek R. Soled, M.D., M.B.A., M.Sc.,[2] and Raja-Elie E. Abdulnour, M.D.[3]

Received: October 28, 2024; Revised: February 3, 2025; Accepted: February 17, 2025; Published: April 24, 2025

Abstract

The rapid integration of artificial intelligence (AI) into health care presents regulatory challenges due to the dynamic, opaque, and adaptive nature of AI — particularly with generative AI models. The U.S. Food and Drug Administration has proposed a life-cycle approach extending beyond conventional frameworks, recognizing that traditional medical device regulations are ill-suited for such technologies. This perspective draws parallels between AI regulation and competency-based medical education (CBME), an existing framework that addresses similar complexities in training human clinicians who are also dynamic general-purpose problem-solvers with opaque cognitive processes. We propose adopting an AI–CBME life-cycle framework for medical AI regulation, leveraging the five core CBME components: defining competencies, sequenced progression, tailored learning experiences, competency-focused instruction, and programmatic assessment. By applying these principles, we can implement continuous outcomes-based assessments of AI systems within their operational environments. This approach embeds real-time validation and safeguards for patient safety, building accountability and trust, while maximizing the potential of AI technologies to enhance health care. Engaging medical educators in this process offers an immediate and practical pathway to develop a robust regulatory framework aligned with existing educational methodologies.

Introduction

The rapid rise of artificial intelligence (AI) in health care demands a dynamic and adaptive regulatory approach. The U.S. Food and Drug Administration (FDA) agrees,[1] but acknowledges that the "traditional paradigm of medical device regulation was not designed for adaptive artificial intelligence and machine learning technologies."[2] The FDA proposes a life-cycle approach to regulating AI that reaches beyond the FDA, spanning the consumer and health care ecosystems, evaluating AI in the settings in which it is deployed, and keeping pace with the rapid progress of AI.[1] The limitations of the current regulatory system are especially evident with the introduction of generative AI models that evolve frequently, can be applied to virtually any clinical problem, and reach conclusions without clear explainability.[3] These models pose unprecedented challenges due to their nondeterministic, dynamic, general-purpose, opaque, and unexplainable nature.[4]

Medical educators encounter similar regulatory challenges daily. Human clinicians are general-purpose, stochastic, and error-prone problem-solving agents whose cognitive processes are largely unexplainable and change continuously with daily and lifelong learning and/or

The author affiliations are listed at the end of the article.

Dr. Abdulnour can be contacted at rabdulnour@bwh.harvard.edu.

disuse. In the early 19th century, educators recognized that a one-size-fits-all medical education did not address this individuality, which led to significant inconsistencies in care.[5] In response, educators adopted a competency-based medical education (CBME) framework emphasizing continuous outcome-based assessment of professional activities for training and certification.[6,7] At its core, CBME was designed to allow flexibility and individualization for students and schools to achieve standardized competencies, which would, in turn, increase patient safety and the quality of care.

The origins of CBME offer a compelling analogy to the need for flexible, outcome-based regulation of medical AI. At first glance, CBME is similar to the FDA's conventional regulatory process (Fig. 1). Like the FDA's premarket phases, medical education begins with preclinical pretraining in knowledge fundamentals with limited and supervised patient exposure. Next, like the postmarket phase, graduate and postlicensure medical education involve life-long education, assessment, and certification in the workplace for the entire life cycle. Progress from one stage to the next depends on assessing competencies and outcomes.

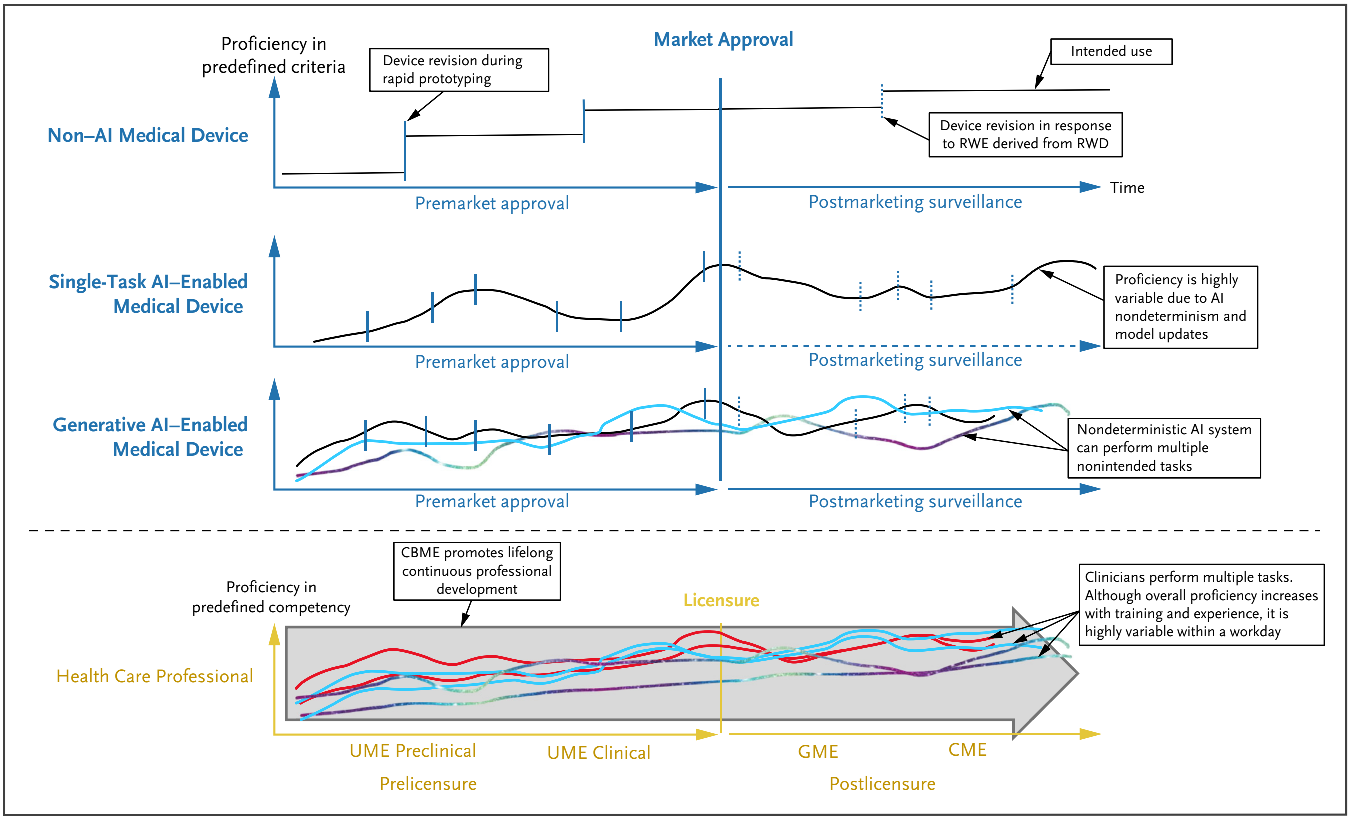

Figure 1. Comparison of Competency-Based Medical Education and the U.S. Food and Drug Administration's Conventional Regulatory Process.

The development of medical devices and the education of health care professionals follow a sequenced approach aimed at promoting training and certification toward defined criteria and competencies. Non-artificial intelligence (AI) medical devices are developed for an intended purpose with minimal performance variability. During premarket approval, rapid prototyping refines the device's accuracy and safety until regulatory approval is granted once predefined criteria are met. Postmarketing collection of real-world data generates real-world evidence that informs further improvements, with proficiency expected to increase over time. Stochastic agents (AI and health care professionals) have variable proficiency learning curves that necessitate evaluation during their entire life cycle, including postauthorization. AI-enabled medical devices are nondeterministic and evolve continuously, requiring more frequent iterations before and after market approval. Generative AI devices add complexity by performing multiple, often unintended tasks, making both premarket evaluation and postmarket surveillance more challenging. Medical learners develop proficiency across various defined competencies and contexts. Competency-based medical education supports lifelong learning through continuous assessment, guiding both professional development and certification for independent clinical practice. AI denotes artificial intelligence; CBME, competency-based medical education; CME, continuing medical education; GME, graduate medical education; RWD, real-world data; RWE, real-world evidence; and UME, undergraduate medical education.

Table 1. Core Principles of Competency-Based Medical Education (CBME) Applied to Human Medical Learners Can Be Adapted to Artificial Intelligence (AI) Systems in Health Care as AI–CBME.*

Core Component	Defining Competencies	Sequence Progression	Tailored Learning Experiences	Competency-Focused Instruction	Programmatic Assessment
CBME example (Van Melle et al.[8])	Medical students must demonstrate proficiency in taking patient histories and diagnosing common conditions	Residents progress from supervised procedures to independently managing patients as they gain skills	Didactics and deliberate practice opportunities designed to broaden a learner's knowledge and skills in different domains are tailored to their needs	Educators design an individualized educational plan based on a learner's abilities in specific tasks	Educators continuously gather learner data and observations from across the workplace and rate them using an entrustment scale. These ratings are passed on to programs and certification bodies to derive assessments against predefined developmental standards, which allow them to determine and document a learner's progression toward independent clinical practice
Core Component	Specific Intended Use	Sequenced Performance Criteria	Trial Phases	Predetermined Change Control Plan	Market Approval and Postmarket Surveillance
AI–CBME example	AI for diagnosing diabetic retinopathy must meet a 95% accuracy threshold compared with human experts	An AI designed to predict sepsis in hospitalized patients must meet increasingly stringent efficacy and safety performance criteria as it progresses from simulated environments using historical patient data, to supervised testing in small cohorts of patients, followed by a randomized controlled clinical trial and deployment	AI systems that predict hospital readmission rates in patients with cardiovascular diseases undergo rapid prototyping and iterative refinements based on trial outcomes in diverse health care settings	An AI system developed to predict hospital readmission rates for patients with cardiovascular diseases is designed to be rapidly modified for different hospital settings, as locally obtained insights inform retraining and fine-tuning	The FDA reviews the results of preclinical and human clinical testing to authorize the marketing of an AI emergency department triage system. During postmarketing surveillance, superusers of the AI system rate their entrustment of the intended task. The health system's AI governing panel aggregates the ratings, which then delivers their assessments to health system administrators and AI developers to guide deployment and development

*Examples for human learners highlight traditional training pathways, while examples for AI demonstrate how similar principles can guide the development, validation, and ongoing assessment of AI tools in clinical practice. AI denotes artificial intelligence; CBME, competency-based medical education; and FDA, U.S. Food and Drug Administration.

Given these similarities, an AI–CBME life-cycle framework represents an immediate opportunity to apply a well-established regulatory approach to medical AI. Working with medical educators, stakeholders can borrow from the five CBME core components identified by Van Melle et al. to implement continuous outcome-based assessment and implementation (Table 1).[8]

1. Defining competencies (CBME) — specific intended use (AI regulatory process): In CBME, clear competencies are outlined to ensure that medical trainees meet specific, measurable outcomes related to their clinical roles. This concept aligns closely with the FDA's definition of a specific intended use,[9] which recognizes that users of AI-enabled devices want the particular indication to be clearly identified. Indeed, defining clear, outcome-based AI competencies — such as improving diagnostic accuracy or predicting patient outcomes — would ensure that AI tools consistently achieve their defined goals. For example, the outcome competency for an AI system designed to diagnose diabetic retinopathy would be retinal image analysis.
2. Sequence progression (CBME) — sequenced performance criteria (AI regulatory process): CBME emphasizes competency development through milestones defined sequentially as learners advance. The regulatory equivalent would be FDA–identified sequenced performance criteria.[10] These criteria explicitly define acceptable performance for an AI system, ensuring that it meets specific, measurable outcomes during sequenced progression from initial testing on small datasets to comprehensive real-world testing, mitigating the premature use of AI systems in clinical practice. For example, during the life cycle of an AI model designed to predict sepsis in hospitalized patients, it must meet increasingly stringent specific accuracy and safety criteria as it progresses from controlled, simulated environments using historical patient data to supervised testing in small cohorts of patients,

followed by a randomized controlled clinical trial and deployment.

3. Tailored learning experiences (CBME) — trial phases (AI regulatory process): Tailored learning experiences in CBME refer to providing learning opportunities through life experiences based on learner needs, promoting lifelong competency development. This principle can be mirrored in AI regulation by emphasizing rapid prototyping during sequenced trial phases. This continual design-evaluation process ensures that AI systems are iteratively refined throughout their life cycle based on data and insights from preclinical and clinical trials, ensuring that they meet the evolving needs of the population where they are deployed. For instance, an AI system developed to predict hospital readmission rates for patients with cardiovascular diseases could undergo iterative refinements based on trial outcomes in diverse health care settings.

4. Competency-focused instruction (CBME) — predetermined change control plan (AI regulatory process): In CBME, competency-focused instruction refers to individualizing an educational plan based on learners' abilities in specific tasks. This involves designing flexible, individualized pathways that adapt to learner evaluations. This parallels the concept of predetermined change control plans (PCCPs) in AI regulation,[11] which document planned modifications to a device, the methodology for validating those modifications, and their impact assessment. Just as competency-focused instruction anticipates and plans for individual growth opportunities, PCCPs provide a structured yet adaptable road map for ongoing AI development as areas for improvement arise. For instance, an AI system developed to predict hospital readmission rates for patients with cardiovascular diseases could be designed to be rapidly tailored for different hospital settings as locally obtained insights inform retraining and finetuning.

5. Programmatic assessment (CBME) — market approval and postmarket surveillance (AI regulatory process): Programmatic assessment in CBME is a systematic approach to continuously gather learner data and observations from across the workplace, and derive actionable insights against predefined developmental standards.[12] Programmatic assessment allows training programs and certification bodies to determine and document a learner's progression toward independent clinical practice. This approach mirrors the use of trial data to inform market authorization,[13] and the use of real-world data to derive real-world evidence from continuous observations across many contexts during postmarketing surveillance to inform regulatory decisions for further use or deployment.[14] Accordingly, market approval and postmarketing surveillance of AI systems would require continuous observation by designated superusers, who would pass on their observations to the health system's AI governing panel. The panel would then deliver its assessments to health system administrators and AI developers to guide deployment and development, respectively.

To facilitate continuous assessment, CBME is coalescing around an entrustable professional activity (EPA) framework where educators evaluate units of observable workplace behavior across many tasks and contexts using a generalizable entrustment scale.[15] These observations are then integrated by competency committees for feedback and entrustment decisions. For example, a novice provider can be entrusted to perform a lumbar puncture only if an educator is at their side (low entrustment). Over time, educators increase their ratings of observations and entrust a lumbar puncture to the learner if observed from a distance (medium entrustment) and later independently (high entrustment). The EPA framework can be a meaningful representation of AI competencies that all stakeholders can relate to, thus facilitating its implementation.[16] For instance, a large language model–based chatbot may be trusted to provide local procedural protocols but not to give diagnostic bedside support.

An AI–CBME life-cycle framework would address the unique regulatory challenges posed by AI by embedding safeguards for real-time validation and patient safety. Aligning regulation with proven educational frameworks promotes a system of accountability and trust while maximizing the potential of AI to improve health care. We can implement such a framework today by inviting medical educators — who already ensure high-quality health care professional skills — into the room where regulation happens.

Disclosures

Author disclosures are available at ai.nejm.org.

Author Affiliations

[1] Academic Chair for Regulation in Law, Medicine, and Technology, Department of Health Sciences and Technology, ETH Zurich and Faculty of Law, University of Zurich, Zurich, Switzerland

[2] Department of Internal Medicine-Pediatrics, Brigham and Women's Hospital, Boston Children's Hospital, Harvard Medical School, Boston

[3] Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston

References

1. Warraich HJ, Tazbaz T, Califf RM. FDA perspective on the regulation of artificial intelligence in health care and biomedicine. JAMA 2024;333:241-247. DOI: 10.1001/jama.2024.21451.

2. FDA. Artificial intelligence and machine learning in software as a medical device. January 6, 2025 (https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device#regulation).

3. Blumenthal D, Patel B. The regulation of clinical artificial intelligence. NEJM AI 2024;1(8). DOI: 10.1056/AIpc2400545.

4. Coiera E, Fraile-Navarro D. AI as an ecosystem — ensuring generative AI is safe and effective. NEJM AI 2024;1(9). DOI: 10.1056/AIp2400611.

5. Flexner A. Medical education in the United States and Canada. from the Carnegie Foundation for the Advancement of Teaching, Bulletin Number Four, 1910. Bull World Health Organ 2002;80:594-602 (https://www.ncbi.nlm.nih.gov/pubmed/12163926).

6. Ten Cate O, Billett S. Competency-based medical education: origins, perspectives and potentialities. Med Educ 2014;48:325-332. DOI: 10.1111/medu.12355.

7. Ten Cate O. Competency-based postgraduate medical education: past, present and future. GMS J Med Educ 2017;34:Doc69. DOI: 10.3205/zma001146.

8. Van Melle E, Frank JR, Holmboe ES, Dagnone D, Stockley D, Sherbino J. A core components framework for evaluating implementation of competency-based medical education programs. Acad Med 2019;94:1002-1009. DOI: 10.1097/ACM.0000000000002743.

9. FDA. General/specific intended use — guidance for industry. March 13, 2018 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/generalspecific-intended-use-guidance-industry).

10. FDA. Safety and performance based pathway. October 11, 2024 (https://www.fda.gov/medical-devices/premarket-notification-510k/safety-and-performance-based-pathway).

11. FDA. Predetermined change control plans for medical devices. August 22, 2024 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/predetermined-change-control-plans-medical-devices).

12. Schuwirth LWT, Vleuten CPMVD. Current assessment in medical education: programmatic assessment. J Appl Test Technol 2019;20:2-10 (https://jattjournal.net/index.php/atp/article/view/143673).

13. FDA. About FDA product approval. December 29, 2017 (https://www.fda.gov/news-events/approvals-fda-regulated-products/about-fda-product-approval).

14. FDA. Real-world evidence. September 19, 2024 (https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence).

15. Ten Cate O, Taylor DR. The recommended description of an entrustable professional activity: AMEE Guide No. 140. Med Teach 2021;43:1106-1114. DOI: 10.1080/0142159X.2020.1838465.

16. Gin BC, O'Sullivan PS, Hauer KE, et al. Entrustment: a framework to safeguard the use of artificial intelligence in health professions education. Acad Med 2025;100:264-272. DOI: 10.1097/ACM.0000000000005930.