International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Brazil's ENAMED and the Governance of AI-Supported Exam Preparation: A Conceptual Policy Analysis

Vinicius Cogo Destefani[1], Matheus Feliciano da Costa Ferreira[2], Afranio Cogo Destefani[3]

[1,2]Centro Universitario Integrado, Campo Mourao, Parana, Brazil.
[3]Escola Superior de Ciencias da Santa Casa de Misericordia de Vitoria (EMESCAM), Vitoria, Espirito Santo, Brazil.

Abstract
Brazil's National Examination for the Assessment of Medical Training (ENAMED) creates a policy setting in which a national assessment can influence student preparation, institutional priorities, and access to residency selection. At the same time, generative artificial intelligence and adaptive learning tools are becoming readily available to medical students. This conceptual policy analysis asks how medical schools should govern AI-supported preparation before high-stakes incentives define its educational purpose. We use a targeted narrative analysis of official ENAMED documents and selected literature on high-stakes testing, medical assessment, artificial intelligence in medical education, and faculty development. We argue that personalisation can become adaptive test coaching when systems optimise visible examination performance without evidence of transfer to clinical reasoning, feedback literacy, or professional judgement. Governance should therefore address educational purpose, evidence standards, transparency, human oversight, equity, data protection, conflicts of interest, and continuous monitoring. Medical schools should align AI-supported preparation with programmatic assessment, prepare faculty and learners to evaluate outputs critically, and keep consequential decisions under accountable human control. ENAMED offers an opportunity to establish these safeguards early. The appropriate question is not whether students will use AI for preparation, but whether institutions will shape that use toward learning and public accountability rather than narrow score optimisation.

Keywords: Medical Education, Artificial Intelligence, Educational Assessment, High-Stakes Testing, Governance, ENAMED, Brazil

1. Introduction

Brazil's Exame Nacional de Avaliacao da Formacao Medica (ENAMED) has placed medical assessment at the intersection of educational quality, public accountability, and progression into residency. Official documents describe ENAMED as a national initiative led by the Ministry of Education and the National Institute for Educational Studies and Research Anisio Teixeira (INEP), with purposes that include evaluating medical training, supporting course improvement, and contributing to selection for direct-entry residency programmes [1] [2]. These purposes make the examination consequential for learners

and institutions even when no single score can represent the full complexity of clinical competence.

The policy change coincides with rapid adoption of generative artificial intelligence, conversational tutoring, automated feedback, and adaptive question systems. These tools can broaden access to practice and explanations. They can also concentrate preparation around what is easiest to measure: recognition of test patterns, response speed, and repeated performance on examination-like items. The central policy issue is therefore not a binary choice between accepting and prohibiting AI. It is whether medical schools will define the educational purpose, acceptable evidence, and safeguards for AI-supported preparation before examination incentives and commercial markets do so for them.

This article develops a conceptual framework for that decision. It does not evaluate an AI product, estimate an effect, or report student outcomes. Its claim is narrower: because high-stakes assessment can reshape teaching and learning, institutions need explicit governance that keeps AI-supported preparation connected to clinical reasoning, professional judgement, and accountable educational practice.

2. Analytical Approach

We conducted a targeted narrative and conceptual analysis rather than a systematic review. The analysis combined official INEP pages and 2025 regulatory materials with purposively selected literature on the effects of high-stakes testing, medical assessment, programmatic assessment, artificial intelligence in medical education, generative AI, and faculty development. Sources were selected to clarify the policy context and to test a chain of reasoning from assessment incentives to institutional governance priorities.

The framework was organised around four questions: what ENAMED makes consequential; how AI-supported personalisation may respond to those incentives; which educational values may be displaced; and which institutional controls can preserve educational purpose. The approach is interpretive. It does not attempt exhaustive literature coverage, formal quality scoring, meta-analysis, or causal inference. Recommendations are therefore presented as governance propositions for local adaptation and evaluation, not as empirically proven effects of a particular intervention.

3. ENAMED as a High-Stakes Policy Context

Assessment influences more than the production of scores. A qualitative metasynthesis of high-stakes testing showed that tests can shape curricular content and instructional form, although effects depend on context [3]. In medical education, assessment is also understood as a programme of judgement that should sample multiple domains, methods, and occasions rather than treating one instrument as a complete representation of competence [4]. Programmatic assessment extends this principle by combining many low- and high-stakes data points with meaningful feedback and proportional decision making [5].

ENAMED's official purposes include evaluating whether graduating students have competencies expected by national curricular guidance, producing information for course improvement, and supporting direct-entry residency selection [1] [2]. These functions create legitimate public value. They also create incentives for students, schools, and preparation providers to focus attention on the examination's most visible outputs. The stronger the perceived consequences of the score, the greater the possibility that preparatory practices will narrow toward test performance.

This is a governance concern, not an argument against national assessment. A standardised examination can inform policy while remaining one part of a broader assessment system. The risk arises when its score becomes a proxy for all educational quality or when local curricula are reorganised primarily

around predictability on the tested format. Medical schools should therefore distinguish preparation that helps learners consolidate knowledge from preparation that displaces workplace assessment, longitudinal feedback, communication, ethics, teamwork, and clinical judgement.

4. From Personalisation to Adaptive Test Coaching

Artificial intelligence has been applied in medical education to tutoring, assessment, prediction, and learning support, but the literature also identifies limitations involving data quality, transparency, implementation, and educator readiness [6] [7]. More recent reviews of large language models describe expanding educational uses alongside risks such as inaccurate output, bias, privacy concerns, and overreliance [8]. Generative AI can support explanation, practice, and formative feedback, yet it does not automatically know which educational goal should be optimised [9].

Personalisation becomes adaptive test coaching when the system's practical objective is reduced to increasing success on examination-like items. That transition can occur gradually. A learner receives questions matched to prior errors; the system favours topics associated with likely score gains; feedback becomes shorter and more answer-centred; and progress dashboards reward item accuracy more visibly than reasoning quality. Each feature may appear useful in isolation, but together they may define learning as prediction of the keyed response.

The distinction should be based on educational function rather than branding. AI-supported preparation is more defensible when it prompts learners to explain reasoning, identifies uncertainty, links answers to reliable sources, encourages reflection, and connects practice to clinical contexts. It becomes more concerning when it presents unsupported certainty, conceals how recommendations are generated, uses sensitive learner data without clear limits, or optimises scores without evidence that learning transfers beyond the test. Claims of effectiveness should specify the outcome measured and should not treat short-term question accuracy as equivalent to clinical competence.

5. Governance Priorities

Medical schools can establish a practical governance framework before AI-supported preparation becomes routine.

First, institutions should define purpose. Approved uses should state whether a tool supports formative practice, feedback, remediation, curriculum mapping, or another objective. The intended benefit, excluded uses, and relationship to existing assessment should be explicit. AI output should not independently determine progression, remediation, or disciplinary decisions.

Second, evidence standards should match the claim. Vendors and internal teams should report more than engagement or score improvement. Evaluation may include reasoning quality, calibration of confidence, feedback usefulness, transfer to unfamiliar cases, accessibility, differential performance across learner groups, and unintended changes in study behaviour. Absence of evidence should be recorded as uncertainty, not converted into a promise of benefit.

Third, transparency and human oversight are essential. Learners should know when content or recommendations are AI-generated, what data are collected, how long data are retained, and how to challenge an error. Faculty must be able to inspect representative outputs and intervene. Consequential judgements require identifiable human responsibility.

Fourth, institutions should address equity and data protection. Unequal access to paid tools may widen preparation gaps, while adaptive systems may reproduce biases present in training data or item banks.

Procurement should examine accessibility, language performance, privacy, security, data reuse, and mechanisms for reporting harm. Commercial relationships and institutional conflicts of interest should be disclosed.

Fifth, monitoring should continue after adoption. Governance is not completed by a one-time approval. Schools should review incidents, learner complaints, output accuracy, participation patterns, equity signals, and evidence of curricular narrowing. A tool should be revised, restricted, or withdrawn when its risks exceed its demonstrated educational value.

6. Implications for Medical Schools

Implementation should connect AI governance to the school's wider assessment strategy. Mapping ENAMED-relevant knowledge to the curriculum may be reasonable, but it should coexist with workplace-based assessment, direct observation, narrative feedback, simulation, and longitudinal review. Programmatic assessment offers a useful organising principle because it resists making one score or one method carry more meaning than it can support [5].

Faculty development is a necessary control, not an optional accompaniment. Educators need opportunities to test prompts, identify hallucinations and bias, judge the quality of feedback, recognise when a learner is outsourcing reasoning, and discuss appropriate disclosure. Early faculty-development reports suggest that structured, hands-on engagement can help medical educators explore generative AI critically, although robust evaluation remains needed [10]. Learners likewise need explicit instruction in verification, attribution, privacy, and the limits of model-generated explanations.

Schools can begin with bounded pilots. A pilot should define users, duration, data flows, success criteria, stop rules, and review responsibility before deployment. It should compare AI-supported preparation with existing educational practice and gather qualitative as well as quantitative signals. Results should be interpreted cautiously: improved performance on similar questions is not evidence of better clinical care. Student representatives, assessment leaders, clinical teachers, privacy officers, and independent reviewers should participate in oversight when the scale or consequences justify it.

7. Limitations

This article is a conceptual policy analysis based on a targeted, non-systematic selection of sources. It may not capture all regulations, implementation changes, or emerging evidence about ENAMED and generative AI. The proposed governance priorities have not been tested as an integrated intervention, and no empirical comparison of preparation tools or student outcomes is presented. Institutional capacity and regulatory obligations vary, so local legal, educational, and privacy review remains necessary. Future research should examine learner behaviour, faculty workload, equity, transfer of learning, and unintended curricular effects using transparent prospective designs.

8. Conclusion

ENAMED makes the governance of AI-supported preparation an immediate institutional responsibility. AI can extend practice and feedback, but personalisation should not be assumed to produce broad educational benefit. Without explicit purpose, proportionate evidence, transparency, human oversight, equity protections, and monitoring, adaptive support may narrow into test coaching. Medical schools should act early: align preparation with programmatic assessment, develop faculty and learner judgement, protect data, disclose conflicts, and keep consequential decisions accountable to people. The

goal is not to prevent students from using AI. It is to ensure that preparation for a high-stakes examination strengthens rather than substitutes for the broader formation of a physician.

Declarations

Conflicts of Interest

Vinicius Cogo Destefani is co-founder and Director of Pedagogy and Engagement at SPRMed (Sao Paulo, Brazil) and a medical advisor at Allm Inc. (Tokyo, Japan). Matheus Feliciano da Costa Ferreira is CEO and Director of Innovation and Artificial Intelligence at SPRMed. Afranio Cogo Destefani has a services agreement with SPRMed and has received consulting fees from the company. Vinicius Cogo Destefani and Afranio Cogo Destefani are family members. SPRMed develops commercial medical-education products, including assessment and learning resources. This article does not evaluate or promote an SPRMed product, dataset, or performance claim. The authors declare no other competing interests.

Ethics Statement

This article reports no research involving human participants, identifiable records, or animals. Ethics committee review and informed consent were not applicable.

Author Contributions

Vinicius Cogo Destefani: conceptualisation, educational policy interpretation, and manuscript review. Matheus Feliciano da Costa Ferreira: conceptualisation, artificial intelligence and governance interpretation, and manuscript review. Afranio Cogo Destefani: conceptualisation, methodology, supervision, writing of the original draft, and review and editing. Final approval by all authors is pending before submission.

Artificial Intelligence-Assisted Technology Disclosure

AI-assisted tools supported structural adaptation, language refinement, reference organisation, and consistency checks for the current editorial version. The authors reviewed the resulting text and remain responsible for all claims, references, disclosures, and the final decision to submit. No AI system generated or analysed study data, made an empirical or clinical judgement, or qualifies for authorship.

Originality Statement

This manuscript is original, has not been published, and is not under consideration by another journal. It reports no participant data.

References

1. Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira. Enamed: 2025 legislation. Brasilia: INEP; 2025. Available from: https://www.gov.br/inep/pt-br/centrais-de-conteudo/legislacao/enamed/2025. Accessed July 17, 2026.
2. Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira. Exame Nacional de Avaliacao da Formacao Medica (Enamed). Brasilia: INEP. Available from: https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enamed/enamed. Accessed July 17, 2026.

3. Au W. High-stakes testing and curricular control: a qualitative metasynthesis. Educational Researcher. 2007;36(5):258-267. doi:10.3102/0013189X07306523.
4. Epstein RM. Assessment in medical education. New England Journal of Medicine. 2007;356(4):387-396. doi:10.1056/NEJMra054784.
5. van der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. A model for programmatic assessment fit for purpose. Medical Teacher. 2012;34(3):205-214. doi:10.3109/0142159X.2012.652239.
6. Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. JMIR Medical Education. 2019;5(1):e13930. doi:10.2196/13930.
7. Gordon M, Daniel M, Ajiboye A, et al. A scoping review of artificial intelligence in medical education: BEME Guide No. 84. Medical Teacher. 2024;46(4):446-470. doi:10.1080/0142159X.2024.2314198.
8. Aster A, Cegla M, Diakowska D, Rosol M. ChatGPT and other large language models in medical education: scoping literature review. Medical Science Educator. 2025;35(1):555-567. doi:10.1007/s40670-024-02206-6.
9. Boscardin CK, Gin B, Golde PB, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. Academic Medicine. 2024;99(1):22-27. doi:10.1097/ACM.0000000000005439.
10. Chadha N, Popil E, Gregory J, Armstrong-Davies L, Justin G. How do we teach generative artificial intelligence to medical educators? Pilot of a faculty development workshop using ChatGPT. Medical Teacher. 2025;47(1):160-162. doi:10.1080/0142159X.2024.2341806.