DEVELOPING AND ASSESSING DIGITAL CLINICAL COMPETENCE THROUGH SIMULATION TASKS AND ELECTRONIC PORTFOLIOS IN INFORMATION TECHNOLOGIES IN MEDICINE

M.M. Rustamov

Associate Professor, Department of Biological Physics,

Informatics and Biomedical Engineering

Andijan State Medical Institute, Andijan, Uzbekistan

Article type: Conceptual article informed by a narrative review

ABSTRACT

Background: Assessment in Information Technologies in Medicine should show whether students can apply digital skills to meaningful medical information tasks. Knowledge tests and self-reported confidence provide incomplete evidence of such performance.

Objective: To propose an assessment framework combining simulated information tasks, an electronic portfolio, and structured performance criteria for undergraduate medical students.

Methods: A focused narrative review of selected informatics education recommendations, electronic health record education studies, and international digital education guidance informed a conceptual assessment design. Targeted searches were conducted on 12 October 2026. No primary student or patient data were collected. Published observations and the framework proposed in this article were kept distinct.

Results: The design comprises four simulated stations addressing information retrieval, data quality, electronic record review, and communication. Five assessment dimensions are proposed: technical accuracy, information reasoning, contextual interpretation, privacy and responsible use, and reflection. A four-level rubric and a portfolio containing original work, feedback, and revision provide complementary evidence. The framework includes a plan for piloting, assessor calibration, and prospective validation.

Conclusion: The proposed system makes digital clinical competence observable through task performance and explanation. Its scores, thresholds, reliability, and relationship to later professional performance remain unvalidated and require empirical investigation.

Keywords: digital clinical competence; medical informatics education; performance assessment; simulation; electronic portfolio; electronic health records; formative feedback.

INTRODUCTION

The educational value of a digital task depends partly on what the student must understand to complete it. Entering information into a form, selecting a graph, and locating a record are observable operations, but their completion alone does not establish that the learner has interpreted the information correctly. Assessment in Information Technologies in Medicine should therefore distinguish successful software interaction from a justified and responsible information decision.

For this article, digital clinical competence is defined operationally as the ability to use digital tools to retrieve, organize, evaluate, and communicate medically contextualized information while explaining relevant limitations and protecting confidentiality. This is an educational working definition. It does not imply that

successful completion of an informatics course authorizes independent clinical practice.

IMIA educational recommendations provide a reference for developing biomedical and health informatics curricula [1]. However, a local assessment design must translate broad educational aims into tasks that can be observed and scored. The objective of this article is to propose such a design through simulation stations and an electronic portfolio, with a clear distinction between the proposed framework and evidence of its effectiveness.

MATERIALS AND METHODS

A focused narrative review was combined with conceptual assessment design. Sources were identified through targeted web searches on 12 October 2026 using phrases concerning medical informatics education, electronic health record simulation, digital assessment, and health workforce educational technology. The selected material comprised IMIA recommendations, primary studies of electronic record education, and WHO and UNESCO publications.

Sources were selected for their direct relevance to task design, evidence of learning, implementation, or responsible use. The search was not exhaustive and did not include independent screening or a formal risk-of-bias assessment. Available publication records and accessible descriptions were used; findings were not pooled quantitatively. The review serves as a rationale for design rather than a systematic estimate of intervention effectiveness.

The design process specified the intended construct, identified situations in which it could be observed, defined the evidence to retain, and proposed scoring criteria. Four task types were developed to sample different aspects of competence. A portfolio was added to capture revision and reflection that a single timed encounter would not show.

No institutional assessment was administered. No actual student marks, group comparisons, reliability coefficients, or patient outcomes were available. The Results section therefore reports the proposed assessment structure. All task descriptions, rubric levels, and validation procedures are illustrative proposals requiring local review and piloting.

RESULTS

Implications of the selected educational evidence

A survey study of medical students' perspectives reported that 88 of 102 respondents considered simulation-based electronic health record training helpful, although students did not feel secure in every aspect of record use [2]. This observation supports distinguishing perceived usefulness from demonstrated competence. It does not establish a causal improvement in professional performance.

An interventional study of electronic health record simulation during an intern boot camp described learner satisfaction and compared navigation with reference workflows. The study did not use a randomized control group [3]. These findings provide a rationale for exploring simulated information tasks, while leaving uncertainty about how performance should be measured in undergraduate informatics courses. Neither study validates the assessment framework proposed here.

Station one and station two

The first proposed station assesses information retrieval. A student receives a short educational question, locates an appropriate source, records a reproducible search strategy, and explains why the source is relevant. The assessor examines the search record and the justification. Naming a familiar website without demonstrating the route to the information is insufficient evidence of the intended outcome.

The second station assesses data quality using synthetic observations. The dataset includes a duplicate record, a missing value, and one inconsistent unit. Students retain the original dataset, document corrections, calculate an appropriate descriptive result, and produce a graph. The assessor observes whether the student distinguishes a data error from an unusual but potentially valid value. The task does not require clinical diagnosis or treatment decisions.

Station three and station four

The third station uses a simulated electronic record. Students identify the correct fictional patient, distinguish historical from current information, locate the date and source of an entry, and describe any unresolved contradiction. The case can be delivered through an educational record interface or a structured mock record. The scoring target is information appraisal rather than familiarity with a particular vendor's menu structure.

The fourth station assesses communication. Students prepare a short handover-style information summary from supplied fictional material and explain which details are necessary for the stated recipient. The assessor examines clarity, traceability, and the proposed sharing method. The student's task is to communicate the provided information accurately, not to issue a real clinical handover or provide patient-specific recommendations.

Common performance dimensions

Five dimensions are proposed across the stations. Technical accuracy concerns whether the digital operations are correct. Information reasoning concerns how sources, transformations, and uncertainties are justified. Contextual interpretation concerns whether the result answers the educational question without exceeding the available evidence. Privacy and responsible use concern appropriate handling and sharing. Reflection concerns the identification and correction of limitations in the student's own work.

Not every dimension needs the same emphasis at every station. Before use, the teacher should produce a blueprint specifying which dimensions are sampled and which actions constitute evidence. This prevents a visually attractive report from receiving credit for skills that were never demonstrated. It also helps ensure that one station does not dominate the course merely because it is easy to score.

Proposed rubric and decision rules

A four-level rubric is proposed for each applicable dimension. Level 0 indicates that the required action is absent or fundamentally incorrect. Level 1 indicates partial completion with substantial prompting or unresolved errors. Level 2 indicates accurate, independent completion of the routine task with an adequate explanation. Level 3 indicates accurate performance accompanied by justified handling of ambiguity or transfer to a modified task.

These levels are descriptors for piloting, not validated proficiency standards. The same numerical label must be supported by station-specific examples. For data quality, for instance, Level 2 may require identifying the inserted errors and documenting corrections; Level 3 additionally requires explaining why automatic deletion of every unusual observation would be inappropriate.

No universal pass mark or weighting is specified. A local panel should decide these only after reviewing the course outcomes, task difficulty, and consequences of the decision. Serious breaches of the simulated confidentiality task should be recorded separately and trigger feedback and reassessment under a published local rule. A high aggregate score should not automatically conceal a material weakness in responsible handling.

Electronic portfolio and feedback

The proposed portfolio contains a limited set of evidence: an initial task response, the submitted work product, assessor feedback, a revised version, and a concise reflective note. The reflection identifies a specific error or uncertainty,

describes the correction, and explains how the student would recognize a similar issue in another task. Unsubstantiated statements such as increased confidence do not replace evidence of improvement.

Portfolio work provides longitudinal information, while the supervised stations provide evidence of independent performance. A short oral explanation can be used to clarify authorship and reasoning when needed. The student should be told in advance which resources are permitted and whether generative artificial intelligence assistance must be declared. The assessment should remain focused on the stated learning outcome rather than on undisclosed monitoring.

DISCUSSION

The proposed framework separates what a learner can produce from how the learner understands the product. A correct answer may arise from genuine competence, close imitation, or unacknowledged assistance. Requiring a documented process and a brief explanation is intended to make those possibilities more distinguishable. It does not eliminate every uncertainty, and it should not be described as a validated method of detecting misconduct.

A key design choice is the use of synthetic information. This permits deliberate insertion of inconsistencies, standardization across students, and repetition without exposing patient data. The educational realism should come from the structure of the information problem and the decisions it requires. It need not depend on access to a live hospital system.

Feasibility is as important as conceptual completeness. WHO's educational technology assessment work highlights the importance of safe, scalable, and cost-conscious implementation [4]. The first pilot could therefore use two stations and a small portfolio rather than a complex platform. Preparation time, assessor workload,

technical failures, and accessibility should be documented together with performance scores.

Generative artificial intelligence creates an additional assessment issue because it can help produce polished work without revealing the learner's reasoning. UNESCO guidance emphasizes human agency and responsible educational use [5]. Within the proposed framework, permitted assistance should be recorded, and any generated factual claim should be checked by the student. An independent supervised task provides a complementary assessment opportunity; an AI-use declaration alone is not evidence of competence.

Prospective validation should begin with expert review of the task blueprint and rubric. Pilot students should then explain how they interpret the instructions, allowing ambiguous wording and unnecessary software demands to be revised. Assessors should independently score sample products, discuss disagreements, and refine anchors before consequential grading begins.

A subsequent study should estimate inter-rater agreement using a statistic appropriate to the score structure and report uncertainty. Performance should be sampled across more than one task, since success on a single spreadsheet exercise may not generalize to information retrieval or record review. Relationships with delayed transfer tasks would provide additional evidence. Self-reported confidence can be recorded, but should be analysed separately from demonstrated performance.

The framework also requires fairness review. Device speed, language, accessibility needs, and previous software exposure can influence observed performance. Equivalent access arrangements and clear instructions should be established before testing. Where time pressure is not part of the intended competence, speed should not become an implicit scoring criterion.

This article has several limitations. The literature selection was purposive, the reviewed studies involve different learner populations, and no local validation has been performed. The proposed rubric does not yet support claims about readiness for clinical work. WHO's research agenda for digitalized health workforce education reinforces the need for further evaluation across contexts [6]. A future study should examine feasibility before testing educational effectiveness.

CONCLUSION

Simulation tasks and an electronic portfolio offer complementary ways to make digital information work visible in Information Technologies in Medicine. The proposed framework specifies task types, performance dimensions, feedback evidence, and a validation pathway. Its principal contribution is an assessable design; its reliability, fairness, and educational value remain questions for empirical research.

REFERENCES

1. Bichel-Findlay J, Koch S, Mantas J, et al. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics: Second Revision. International Journal of Medical Informatics. 2023;170:104908. https://doi.org/10.1016/j.ijmedinf.2022.104908

2. Herrmann-Werner A, Holderried M, Loda T, Malek N, Zipfel S, Holderried F. Navigating Through Electronic Health Records: Survey Study on Medical Students' Perspectives in General and With Regard to a Specific Training. JMIR Medical Informatics. 2019;7(4):e12648. https://doi.org/10.2196/12648

3. Miller ME, Scholl G, Corby S, Mohan V, Gold JA. The Impact of Electronic Health Record–Based Simulation During Intern Boot Camp: Interventional Study. JMIR Medical Education. 2021;7(1):e25828. https://doi.org/10.2196/25828

4. World Health Organization. Health workforce education technology assessment: background research report. Geneva: WHO; 2023. https://www.who.int/publications/i/item/9789240070929

5. UNESCO. Guidance for generative AI in education and research. Paris: UNESCO; 2023. https://unesdoc.unesco.org/ark:/48223/pf0000386693

6. World Health Organization. Digitalized health workforce education: an elicitation of research gaps and selection of case studies. Geneva: WHO; 2023. https://www.who.int/publications/i/item/9789240072718