Research Article

From Isolated Exams to Longitudinal Judgment: Implementing Programmatic Assessment with Multisource Data and Portfolio-Based Educational Decisions in Anesthesiology Residency

Firoozeh Madadi [iD] [1], Amir Maziar [2], Reza Shahghadami [3], Hossein Bonakchi [iD] [4], Ali Dabbagh [iD] [1, *]

[1] Anesthesiology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2] Educational Development Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3] Department of Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4] Department of Epidemiology and Biostatistics, School of Health, Torbat Heydariyeh University of Medical Sciences, Tehran, Iran

Corresponding Author: Department of Anesthesiology, Anesthesiology Research Center, Shahid Modarres Hospital, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: alidabbagh@yahoo.com

Abstract

Background: Competency-based medical education (CBME) emphasizes longitudinal development of professional competence and has driven a shift away from isolated high-stakes examinations toward programmatic assessment (PA). Programmatic assessment integrates multiple low-stakes assessments, continuous feedback, and portfolio-based decision-making to support learning and ensure defensible judgments of competence. Evidence regarding the implementation of PA in anesthesiology residency remains limited.

Objectives: The main objective of this study was to assess the efficiency of PA in anesthesiology residency in Iran for improving the quality of anesthesiology training.

Methods: This mixed-methods descriptive-analytic study evaluated the implementation of a PA system in an anesthesiology residency program at a university-affiliated teaching hospital in Iran (2020 - 2025). The assessment framework incorporated entrustable professional activities (EPAs), workplace-based assessments [mini-clinical evaluation exercise (mini-CEX) and direct observation of procedural skills (DOPS)], bi-monthly multiple-choice question (MCQ) examinations with individualized feedback, multisource feedback, global faculty assessments, and an electronic portfolio. Longitudinal resident performance data were analyzed using descriptive statistics and linear mixed-effects models. Faculty mentors' attitudes were assessed using a validated questionnaire.

Results: Sixty-one residents contributed 134 resident-year observations across four years of training. Mean scores for direct observational performance assessments, EPAs, and global faculty assessments increased progressively with advancing training year. Mixed-effects analyses demonstrated a significant effect of training year on all outcomes (P < 0.001). Faculty mentors reported positive attitudes toward the PA system, high satisfaction, and strong willingness to continue participation, indicating acceptability and perceived educational value.

Conclusions: Programmatic assessment was feasible to implement in anesthesiology residency and was associated with longitudinal improvement in resident performance across competency domains. Faculty acceptance further supports the educational value of PA as an effective strategy for operationalizing CBME and supporting evidence-informed, learner-centered educational decision-making.

Keywords: Programmatic Assessment, Anesthesiology Residency, Competency Based Medical Education, Entrustable Professional Activities, Direct Observation of Procedural Skills

1. Introduction

Competency-based medical education (CBME) is one of the latest transformational approaches in medical education, which has revolutionized residency training by stressing the role of "progressive development in professional competence" compared with the traditional time-based approach, which considers fixed intervals as the core of educational schools ([1, 2]). This novel paradigm surpasses the traditional high-stakes testing. The underlying logic is clear: One-time examinations are increasingly viewed as insufficient for evaluating the meticulous realities of clinical practice, especially in challenging, high-stakes fields like anesthesiology. Accordingly, revised assessment strategies have gained increasing appeal so we can focus on a trend of "continuing data collection", "timely and meaningful feedback", and "conscientious, evidence-based decisions" regarding the trajectory of trainee progress and development ([3, 4]). Programmatic

assessment (PA) provides a comprehensive and all-inclusive structure that incorporates several low-stakes assessment tools in a goal-directed, determined, longitudinal manner, yielding comprehensive insights into a trainee's competence (5, 6). Instead of relying on isolated testing incidents, it embraces "workplace-based observations", "multisource feedback", "reflective portfolios", and "objective metrics", with feedback providing the main impetus for development (7). Decisions about educational advancements, remediation, or graduation are reached through careful gathering and aggregation of these varied-source data and their logical and thoughtful interpretation, typically within a structured portfolio system (1, 2, 8).

In anesthesiology residency, where patient safety is highly dependent on technical expertise, non-technical skills, and instantaneous clinical judgment and discernment, PA presents a promising, educationally powerful substitute to traditional methods. Nevertheless, real-world adoption faces barriers, including the need for additional faculty support, effective data management, and establishing the trustworthiness and acceptance of portfolio-driven decisions (9, 10).

This study describes and evaluates the introduction of a PA system in an anesthesiology residency program in the Department of Anesthesiology, School of Medicine, Shahid Beheshti University of Medical Sciences (DACCPM, SBMU), Tehran, Iran, with specific focus on the integration of diverse "assessment tools", "standardized feedback mechanisms", "multisource data collection", and "justifiable, portfolio-oriented decision-making processes".

2. Objective

The primary aim of the study is to provide empirical evidence regarding the practicality and educational benefits of PA in postgraduate anesthesiology training.

3. Methods

In this mixed-methods descriptive-analytic study, we aimed to describe and evaluate the implementation of PA in the anesthesiology residency program, in the DACCPM, SBMU, Tehran, Iran. The study focused on the design and use of multisource assessments, the integration of knowledge and clinical practice data, the role of educational feedback, and evidence-based training decision-making in the form of an educational portfolio.

3.1. Ethical Considerations

This study was conducted after obtaining permission from the Research Ethics Committee, Deputy of Research, SBMU, Tehran, Iran (ethics approval code: IR.SBMU.RETECH.REC.1404.728). Participation of all individuals was voluntary, and data confidentiality and anonymity of respondents were fully respected.

3.2. Study Setting and Participants

The study was conducted in the DACCPM, SBMU, Tehran, Iran, during a 5-year period from 2020 - 2025. The research population included all anesthesiology residents who were in training during their residency course and also the faculty members of the department.

3.3. Programmatic Assessment Framework and Data Sources

The PA model was designed based on the principles of competency-based education and emphasized the continuous collection of low-risk, longitudinal data from a variety of sources. In this framework, no single evaluation tool was the sole basis for educational decision-making, and triangulation of evidence was a fundamental principle of educational judgment.

3.4. Data Collection Methods

The data sources used included the following (summarized in Table 1):

1. Entrustable professional activities (EPAs): Entrustable professional activities were designed indigenously and in-house (by the faculty members of DACCPM, SBMU) and developed based on the national anesthesiology residency curriculum; the process of EPA development has been described in another study (11). In total, 40 EPAs covered different areas of clinical practice, decision-making, technical and non-technical skills, and anesthesia patient management. These EPAs were assessed frequently and in real-world settings by faculty, and their results were recorded.

2. Multiple-choice written tests: Multiple-choice questions (MCQs) were administered bi-monthly and went beyond simply measuring knowledge. Each exam included 150 standard MCQs and was held electronically. The test results were provided to the residents along with direct individual feedback that included

Table 1. Core Components of the Programmatic Assessment System Mapped to ACGME Competency Framework

Component	Description	Educational Purpose [a]
EPAs	Forty in-house – developed EPAs covering clinical decision-making, technical and non-technical skills, and anesthesia patient management	Assessment of Patient care, medical knowledge, and interpersonal and communication skills through real-world entrustment decisions
Workplace-based assessments (mini-CEX, DOPS)	Direct observation of clinical encounters and procedural skills with immediate structured feedback	Direct evaluation of Patient care, procedural skills, and professionalism in authentic clinical settings
MCQ examinations	Bi-monthly MCQ tests with individualized feedback and gap analysis	Monitoring progression in medical knowledge and supporting evidence-based learning
Multisource feedback	Feedback from faculty, peers, mentors, and resident self-assessments	Comprehensive assessment of professionalism, interpersonal and communication skills, and practice-based learning and improvement
Global faculty assessment	Longitudinal professional judgment of overall performance, accountability, and consistency	Holistic evaluation of professionalism, patient care, and systems-based practice
Educational portfolio	Longitudinal aggregation of multisource assessment data within an electronic EPA-embedded logbook	Integration of evidence across all six ACGME competencies to support longitudinal judgment
Portfolio-based competency committee decisions	Structured committee review of aggregated data for progression or remediation	Competency-based decision-making aligned with ACGME milestones and practice-based learning and improvement

Abbreviations: EPAs, entrustable professional activities; mini-CEX, mini-clinical evaluation exercise; DOPS, direct observation of procedural skills; MCQ, multiple-choice question.

[a] Educational purposes are explicitly aligned with the six ACGME core competencies: Patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice, ensuring that assessment components support competency-based progression and milestone-driven educational decisions.

performance analysis, identification of knowledge gaps, and provision of targeted educational content tailored to identified needs (12-15). Data from MCQs were considered a key pillar of educational decision-making and played a critical role in identifying the need for educational support, planning remedial interventions, and assessing academic progress (14).

3. Direct observational performance assessments: Included the mini-clinical evaluation exercise (mini-CEX) to assess clinical performance in real-world situations and the direct observation of procedural skills (DOPS) to measure procedural skills. These assessments were administered in a structured manner and provided immediate feedback (12, 13, 16).

4. Supplementary educational data: Included feedback from mentor faculty, resident self-assessments, and interactive reflections recorded in portfolios.

5. Global assessment: In addition to the structured instruments, the faculty global assessment was considered a complementary source of data. It reflected faculty professional judgment about the overall performance of residents over specified time periods and covered aspects such as professional accountability, clinical judgment, communication, and consistency of performance in real clinical situations. The global assessment was not used alone as the basis for educational decision-making, but was used in conjunction with other multisource data as a tool to

complete the comprehensive picture of residents' performance.

6. Electronic EPA-embedded logbook: A custom-designed intra-departmental electronic logbook with EPAs embedded was used to systematically and seamlessly record assessment data. This electronic logbook allowed for the recording of EPAs, mini-CEX and DOPS assessments, structured faculty feedback, and resident feedback. The use of this electronic system facilitated longitudinal data collection, increased information accessibility for mentors, and enhanced transparency in the portfolio-based decision-making process of anesthesiology residents.

3.5. Educational Portfolio and Data Integration

All data from EPAs, MCQs with educational feedback, mini-CEX, and DOPS were aggregated into a structured educational portfolio for each resident, based on previous experiences (12). All data from EPAs, MCQ tests with educational feedback, direct observation assessments (mini-CEX and DOPS), faculty global assessments (global assessment), and other educational documentation were integrated into each resident's training portfolio through an electronic logbook embedded with EPA. The portfolio was the primary data integration tool, allowing for longitudinal review of performance, distinguishing knowledge problems from

Table 2. Targeted Educational Interventions for Improving Residents' Academic and Clinical Performance Within the Programmatic Assessment Framework

Identified Performance Pattern	Underlying Issue	Targeted Intervention	Educational Outcome
Adequate knowledge with weak clinical performance	Performance anxiety, situational stress, or difficulty applying knowledge in real settings	Mentoring support, structured feedback, counseling referral, increased supervised clinical exposure	Improvement in clinical performance and decision-making
Strong clinical skills with poor examination performance	Ineffective study strategies or test-taking skills	Individualized remedial education, focused MCQ practice, feedback-driven learning plans	Progressive improvement in knowledge-based assessments
Slow overall progress across competencies	Learning gaps or inconsistent engagement	Early identification through portfolio review and personalized learning plans	Prevention of academic failure and timely support
Discrepancy between assessment tools	Overreliance on single assessment method	Triangulation of multisource data and holistic judgment	Fairer and more defensible competency decisions
High workload or feedback fatigue	Faculty or resident burden	Streamlined electronic logbook and structured feedback templates	Improved feasibility and sustainability of assessment

Abbreviation: MCQ, multiple-choice question.

performance weaknesses, and identifying patterns of educational progress or decline.

3.6. Evidence-Based Decision Making in Anesthesiology Residents' Education

Education-related decision making, including continuing pathways in anesthesiology residents' medical education, the need for targeted support or targeted continuing education, and the design of remedial interventions, were based on a consolidated review of multisource data contained in the portfolio (Table 2). In this process, MCQ test scores were interpreted proactively and in interaction with the results of EPAs and performance assessments, and no single instrument was the sole basis for instructional judgment (12, 13).

3.7. Faculty Mentor Experience Assessment Tool

A structured questionnaire consisting of 8 questions was designed to assess faculty mentors' experience and satisfaction with the implementation of the study evaluation model. The questionnaire covered dimensions such as the adequacy of multisource data, the possibility of providing meaningful feedback, the role of knowledge data in educational decision-making, the validity of portfolio-based judgments, mentors' workload, and overall satisfaction with participating in program evaluation. Their responses were recorded on a five-point Likert scale with the following response order: Strongly disagree = 1, disagree = 2, no opinion = 3, agree = 4, and strongly agree = 5.

3.8. Validity and Reliability of the Assessment Tool

The content validity of the questionnaire was reviewed and revised by consulting experts in medical education and anesthesia. The internal validity of the assessment tool was increased using Cronbach's alpha coefficient.

4. Results

A total of 61 residents contributed 134 resident-year observations across four years of training. Descriptive statistics demonstrated progressive increases in all assessment metrics over residency years (Table 1).

4.1. Descriptive Longitudinal Trends in Resident Performance

The main outcome metrics were: Mini-clinical evaluation exercise global rating score, range 0 - 10; DOPS holistic-level score, range 0 - 10; EPAs, range 1 - 5; and MCQ bi-monthly examination score, out of 150.

Linear mixed-effects models accounting for repeated measurements within residents showed a significant effect of training year on all outcomes. Compared with year 1, mean mini-CEX scores increased significantly in years 2, 3, and 4 (all $P < 0.001$; Table 3). Similar patterns were observed for DOPS scores, with large and statistically significant increases across successive training years. MCQ scores also increased substantially with advancing training year, with residents in year 4 demonstrating nearly a 19-point higher mean MCQ score compared with year 1 ($P < 0.001$).

Workplace-based assessment metrics were strongly associated with monthly global assessment scores. In mixed-effects models adjusted for training year, both direct observational performance assessments and EPA independently demonstrated significant positive associations with global assessment scores, indicating

Table 3. Resident-Year Outcomes Across Residency Training [a]

Training	Residents (N)	Direct Observational Performance Assessment	EPA	Global Assessment
Year 1	61	6.17 ± 0.55	1.92 ± 0.09	119.31 ± 0.89
Year 2	37	6.82 ± 0.60	3.75 ± 0.48	129.46 ± 3.46
Year 3	22	7.50 ± 0.57	4.55 ± 0.21	132.10 ± 4.64
Year 4	14	7.70 ± 0.90	4.68 ± 0.33	135.19 ± 7.27

Abbreviation: EPA, entrustable professional activity.
[a] Values are expressed as mean ± standard deviation.

that higher longitudinal workplace-based performance was associated with better performance in educational rotations. Due to the high correlation between observational assessment tools and EPAs, a combined model including both predictors demonstrated convergence instability and was not retained.

All three metrics show a monotonic increase across training years, consistent with progressive competency development.

4.2. Attitude Assessment Results

The reliability of the attitude assessment tools for both faculty mentors and anesthesiology residents was evaluated using Cronbach's alpha coefficient. The Cronbach's alpha value for the faculty mentors' attitude assessment tool was 0.82, indicating appropriate reliability and desirable internal consistency. Similarly, the anesthesiology residents' attitude assessment tool demonstrated a Cronbach's alpha value of 0.85, reflecting good internal consistency and reliability of the instrument. The content validity of both questionnaires was confirmed through careful design of the items based on the objectives of the program evaluation system and validation by expert opinion.

The descriptive results of faculty mentors' attitudes are presented in Table 4. As shown, the mean scores for all items were above the average level, indicating an overall positive attitude toward the evaluation system. The highest mean score was related to the willingness to continue participating as a mentor, reflecting a high level of acceptance and satisfaction among mentors. Moreover, lower standard deviations observed in certain items, such as the ability to provide meaningful feedback, suggest convergence in mentors' perspectives.

The descriptive results of anesthesiology residents' attitudes are presented in Table 5. The mean scores for all items were also above average, demonstrating

generally favorable perceptions of the evaluation system among residents. The highest mean scores were observed for the items "Effect of Frequent Low-Risk Evaluations on Learning" and "The Role of Mentoring in Increasing Active Participation", indicating good acceptance of the system and satisfaction with its educational impact.

5. Discussion

In this study, for the first time, a PA system based on CBME principles was systematically designed, implemented, and evaluated in the anesthesia residency program of SBMU, Tehran, Iran. This novel approach, emphasizing the continuous collection of low-risk data from diverse sources, providing regular and timely feedback, and educational decision-making based on an electronic portfolio, provided a suitable alternative to traditional high-risk, single-point assessments (12-14). The results obtained indicated high feasibility, operational sustainability, and positive impact of this system on the development and advancement of the residents' competencies during the four-year training period.

The main findings of the study confirmed a significant and consistent increase in scores on all assessment instruments. Scores on the mini-CEX, DOPS, EPAs, and MCQs increased significantly with advancing years of residency (P < 0.001). This unidirectional and monotonic pattern is fully consistent with the theoretical foundations of CBME and indicates a gradual and natural development of competencies in the areas of patient care, medical knowledge, professionalism, and the other six ACGME competencies (17). In addition, strong correlations were observed between workplace-based assessments (direct observation and EPAs) and faculty global assessments (global assessment), confirming the internal validity and consistency of the system and indicating that multisource data can

Table 4. Faculty Mentors' Attitudes Toward the Programmatic Assessment System (N = 25) [a]

Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean ± Standard Deviation
The statement in the questionnaire	0 (0)	2 (8)	5 (20)	14 (56)	4 (16)	3.68 ± 0.99
Clarity of the mentor's role	0 (0)	4 (16)	7 (28)	9 (36)	5 (20)	3.20 ± 1.15
Adequacy of assessment tools	0 (0)	1 (4)	3 (12)	17 (68)	4 (16)	3.96 ± 0.68
Ability to provide meaningful feedback	0 (0)	3 (12)	5 (20)	7 (28)	10 (40)	3.96 ± 1.06
Value of multi-source data in assessment	0 (0)	2 (8)	5 (20)	14 (56)	4 (16)	3.80 ± 0.82
Validity of instructional decisions	0 (0)	4 (16)	3 (12)	11 (44)	7 (28)	3.84 ± 1.03
Workload and feasibility of implementation	0 (0)	3 (12)	5 (20)	9 (36)	8 (32)	3.88 ± 1.01
Overall satisfaction with mentor role	0 (0)	0 (0)	2 (8)	11 (44)	12 (48)	4.40 ± 0.65

[a] Values are expresses as No. (%) unless otherwise indicated.

Table 5. Anesthesiology Residents' Attitudes Toward the Programmatic Assessment System (N = 56) [a]

Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean ± Standard Deviation
Clarity of the overall structure of the programmatic assessment system	5 (8.0)	10 (16.1)	19 (30.6)	17 (27.4)	11 (17.7)	2.88 ± 1.22
Feedback helped identify strengths and weaknesses	2 (3.2)	16 (25.8)	21 (33.8)	22 (35.4)	1 (1.6)	2.98 ± 0.96
Frequent assessments improved continuous and goal-oriented learning	4 (6.4)	13 (20.9)	16 (25.8)	22 (35.4)	7 (11.2)	3.13 ± 1.15
Mentoring and reflection increased active engagement in learning	7 (11.2)	11 (17.7)	19 (30.6)	18 (29.0)	7 (11.2)	3.13 ± 1.15
Frequent assessments enhanced continuous learning	4 (6.4)	16 (25.8)	18 (29.0)	20 (32.2)	4 (6.4)	3.00 ± 1.06
Overall improvement in educational quality and professional readiness	3 (4.8)	13 (20.9)	18 (29.0)	23 (37.0)	5 (8.0)	3.09 ± 1.10
Overall contribution of mentoring to academic progress	7 (11.2)	7 (11.2)	20 (32.2)	17 (27.4)	11 (17.7)	2.96 ± 1.13

[a] Values are expresses as No. (%) unless otherwise indicated.

provide a comprehensive, valid, and holistic picture of resident performance.

One of the outstanding strengths of this system was the successful integration of multisource data into a customized electronic portfolio based on the EPA-embedded logbook. This tool not only enabled longitudinal monitoring of performance, but also facilitated early detection of problematic patterns (such as gaps between theoretical knowledge and clinical practice) and the design of targeted and personalized educational interventions ([18]). As described in Table 2, interventions such as increased mentoring, targeted MCQ practice, increased supervised clinical exposure, and individualized learning programs resulted in sustained improvements in performance and prevention of academic failure. This data triangulation approach minimized the risk of making decisions based on a single tool and increased the fairness and defensibility of the eligibility committee's decisions ([19]).

The survey of participants' attitudes also yielded interesting results. Faculty mentors showed a generally positive attitude; the mean scores for most items were above 3.8, and the highest satisfaction was observed in the areas of continued mentorship (4.40) and adequacy of assessment tools. These findings indicate high acceptance among faculty and the perceived value of the system in providing meaningful feedback. In contrast, residents' attitudes were more cautious and ambivalent; the mean score was around 3 and a significant percentage expressed a neutral or negative opinion. However, items related to the impact of frequent assessments on continuous learning, the role of mentoring in increasing active participation, and the overall improvement of educational quality received the highest scores. This difference in attitude is likely due to the additional workload of frequent assessments, initial ambiguity in the structure of the system, or natural resistance to change from the traditional time-based

model. Similar experiences have been reported in international studies, emphasizing that residents' acceptance increases over time and with better training (14, 20).

Comparing our results with similar studies in other countries, such as PA systems in US anesthesia residencies by Woodworth et al. (9) or a mobile app in Switzerland by Marty et al. (10), reveals common implementation challenges such as increased faculty workload and the need for cultural acceptance. Nevertheless, the use of a customized electronic logbook in the present study significantly improved transparency, rapid access to data, and reduced administrative burden, and could be proposed as a practical solution for similar centers in developing countries (21-23).

Despite its strengths, the study had limitations. Implementation at a single academic department, a limited sample of senior residents (due to the cohort nature of the study), and a descriptive-analytic design without a control group limit the generalizability of the results to some extent. Also, the long-term effects of the system on national board exam success, post-graduation clinical performance, or patient safety were not examined. It is suggested that future studies with a prospective cohort design, direct comparison with traditional programs, and long-term follow-up of graduates be conducted to determine the true impact on clinical and professional outcomes.

5.1. Conclusions

Implementing a program evaluation in anesthesia residency in Iran is not only feasible and practical, but also has significant educational benefits and could serve as a model for reforming educational programs in other specialties. Given the dual perspective of residents, it is recommended that in the early stages of implementation, a focus be placed on extensive stakeholder education, streamlining processes, reducing unnecessary assessments, and fostering a culture of continuous and supportive feedback to maximize adoption. This approach is an important and strategic step in aligning Iranian postgraduate medical education with global competency-based standards and has the potential to enhance the quality of anesthesia care and patient safety in the long term. A brief description of the study could be found as an infographic in Figure 1. This infographic has been drawn by ChatGPT 5.2.

Figure 1. A brief description of the study

Footnotes

AI Use Disclosure: ChatGPT 5.2. was used completely to draw Figure 1.

Authors' Contribution: Study concept and design: A. D. and F. M.; Acquisition of data: A. D., F. M., and A. M. N.; Analysis and interpretation of data: A. D., F. M., R. S., and H. B.; Drafting of the manuscript: A. D. and F. M.; Critical revision of the manuscript for important intellectual content: A. D., F. M., A. M. N., R. S., and H. B.; Statistical analysis: A. D., F. M., R. S., and H. B.; Administrative, technical, and material support: A. D. and F. M.; Study supervision: A. D.

Conflict of Interests Statement: The authors declare no conflict of interests. Firoozeh Madadi and Ali Dabbagh are members of the editorial board of this journal; however, they had no role in the peer-review

process or editorial decision-making for this manuscript.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication.

Ethical Approval: This study was conducted after obtaining permission from the Research Ethics Committee, Deputy of Research, Shahid Beheshti University of Medical Sciences, Tehran, Iran (IR.SBMU.RETECH.REC.1404.728).

Funding/Support: This study was supported in part by grant 43018584 from the Deputy of Research, Shahid Beheshti University of Medicine, Tehran, Iran.

Informed Consent: Written informed consent was obtained from the participants.

References

1. Hamza DM, Hauer KE, Oswald A, van Melle E, Ladak Z, Zuna I, et al. Making sense of competency-based medical education (CBME) literary conversations: A BEME scoping review: BEME Guide No. 78. *Med Teach.* 2023;**45**(8):802-15. [PubMed ID: 36668992]. https://doi.org/10.1080/0142159X.2023.2168525.

2. Torre D, Schuwirth L. Programmatic assessment for learning: A programmatically designed assessment for the purpose of learning: AMEE Guide No. 174. *Med Teach.* 2025;**47**(6):918-33. [PubMed ID: 39368061]. https://doi.org/10.1080/0142159X.2024.2409936.

3. Alharbi NS. Evaluating competency-based medical education: a systematized review of current practices. *BMC Med Educ.* 2024;**24**(1):612. [PubMed ID: 38831271]. [PubMed Central ID: PMC11149276]. https://doi.org/10.1186/s12909-024-05609-6.

4. Iobst WF, Holmboe ES. Programmatic Assessment: The Secret Sauce of Effective CBME Implementation. *J Grad Med Educ.* 2020;**12**(4):518-21. [PubMed ID: 32879699]. [PubMed Central ID: PMC7450746]. https://doi.org/10.4300/JGME-D-20-00702.1.

5. Misra S, Iobst WF, Hauer KE, Holmboe ES. The Importance of Competency-Based Programmatic Assessment in Graduate Medical Education. *J Grad Med Educ.* 2021;**13**:113-9. [PubMed ID: 33936544]. [PubMed Central ID: PMC8078068]. https://doi.org/10.4300/JGME-D-20-00856.1.

6. de Jong LH, Bok HGJ, Schellekens LH, Kremer WDJ, Jonker FH, van der Vleuten CPM. Shaping the right conditions in programmatic assessment: how quality of narrative information affects the quality of high-stakes decision-making. *BMC Med Educ.* 2022;**22**(1):409. [PubMed ID: 35643442]. [PubMed Central ID: PMC9148525]. https://doi.org/10.1186/s12909-022-03257-2.

7. Gupta SK, Srivastava T. Assessment in Undergraduate Competency-Based Medical Education: A Systematic Review. *Cureus.* 2024;**16**(4). https://doi.org/10.7759/cureus.58073.

8. Schut S, Maggio LA, Heeneman S, van Tartwijk J, van der Vleuten C, Driessen E. Where the rubber meets the road - An integrative review of programmatic assessment in health care professions education.

Perspect Med Educ. 2021;**10**(1):6-13. [PubMed ID: 33085060]. [PubMed Central ID: PMC7809087]. https://doi.org/10.1007/s40037-020-00625-w.

9. Woodworth GE, Goldstein ZT, Ambardekar AP, Arthur ME, Bailey CF, Booth GJ, et al. Development and Pilot Testing of a Programmatic System for Competency Assessment in US Anesthesiology Residency Training. *Anesth Analg.* 2024;**138**(5):1081-93. [PubMed ID: 37801598]. https://doi.org/10.1213/ANE.0000000000006667.

10. Marty AP, Braun J, Schick C, Zalunardo MP, Spahn DR, Breckwoldt J. A mobile application to facilitate implementation of programmatic assessment in anaesthesia training. *Br J Anaesth.* 2022;**128**(6):990-6. [PubMed ID: 35410792]. https://doi.org/10.1016/j.bja.2022.02.038.

11. Dabbagh A, Fadaeizadeh L, Gharaei B, Ghasemi M, Kamranmanesh M, Khorasanizadeh S, et al. The Role of Entrustable Professional Activities in Competency-based Medical Education for Anesthesiology Residents: A Pilot Phase. *Anesthesiol Pain Med.* 2022;**12**(5). https://doi.org/10.5812/aapm-130176.

12. Dabbagh A, Gandomkar R, Farzanegan B, Jaffari A, Massoudi N, Mirkheshti A, et al. Residency Education Reform Program in Department of Anesthesiology and Critical Care: An Academic Reform Model. *Anesthesiol Pain Med.* 2021;**11**(3). https://doi.org/10.5812/aapm.113606.

13. Dabbagh A, Elyassi H, Sabouri A, Vahidshahi K, Ziaee SAM. The Role of Integrative Educational Intervention Package (Monthly ITE, Mentoring, Mocked OSCE) in Improving Successfulness for Anesthesiology Residents in the National Board Exam. *Anesthesiol Pain Med.* 2020;**10**(2). https://doi.org/10.5812/aapm.98566.

14. Dabbagh A, Massoudi N, Vosoghian M, Mottaghi K, Mirkheshti A, Tajbakhsh A, et al. Improving the Training Process of Anesthesiology Residents Through the Mentorship-Based Approach. *Anesth Pain Med.* 2019;**9**(1). e88657. [PubMed ID: 30881915]. [PubMed Central ID: PMC6412912]. https://doi.org/10.5812/aapm.88657.

15. Sezari P, Tajbakhsh A, Massoudi N, Arhami Dolatabadi A, Tabashi S, Sayyadi S, et al. Evaluation of One-Day Multiple-Choice Question Workshop for Anesthesiology Faculty Members. *Anesthesiol Pain Med.* 2020;**10**(6). https://doi.org/10.5812/aapm.111607.

16. Dabir S, Hoseinzadeh M, Mosaffa F, Hosseini B, Dahi M, Vosoughian M, et al. The Effect of Repeated Direct Observation of Procedural Skills (R-DOPS) Assessment Method on the Clinical Skills of Anesthesiology Residents. *Anesthesiol Pain Med.* 2021;**11**(1). https://doi.org/10.5812/aapm.111074.

17. Lin HJ, Wu JH, Lin WH, Nien KW, Wang HT, Tsai PJ, et al. Using ACGME milestones as a formative assessment for the internal medicine clerkship: a consecutive two-year outcome and follow-up after graduation. *BMC Med Educ.* 2024;**24**(1):238. [PubMed ID: 38443912]. [PubMed Central ID: PMC10916194]. https://doi.org/10.1186/s12909-024-05108-8.

18. Dabbagh A, Madadi F, Larijani B. Role of AI in Competency-Based Medical Education: Using EPA as the Magicbox. *Arch Iran Med.* 2024;**27**(11):633-5. [PubMed ID: 39534999]. [PubMed Central ID: PMC11558609]. https://doi.org/10.34172/aim.31795.

19. Duran HT, Kingeter M, Reale C, Weinger MB, Salwei ME. Decision-making in anesthesiology: will artificial intelligence make intraoperative care safer? *Curr Opin Anaesthesiol.* 2023;**36**(6):691-7. [PubMed ID: 37865848]. [PubMed Central ID: PMC11100504]. https://doi.org/10.1097/ACO.0000000000001318.

20. Joe MB, Cusano A, Leckie J, Czuczman N, Exner K, Yong H, et al. Mentorship Programs in Residency: A Scoping Review. *J Grad Med*

Educ. 2023;**15**(2):190-200. [PubMed ID: 37139208]. [PubMed Central ID: PMC10150829]. https://doi.org/10.4300/JGME-D-22-00415.1.

21. Barbieri A, Giuliani E, Lazzerotti S, Villani M, Farinetti A. Education in anesthesia: three years of online logbook implementation in an Italian school. *BMC Med Educ.* 2015;**15**:14. [PubMed ID: 25881277]. [PubMed Central ID: PMC4331334]. https://doi.org/10.1186/s12909-015-0298-1.

22. Sehmbi H, Shah UJ. Electronic logbooks for residents: A step forward. *Indian J Anaesth.* 2013;**57**(2):210-2. [PubMed ID: 23825833]. [PubMed Central ID: PMC3696281]. https://doi.org/10.4103/0019-5049.111878.

23. McGinn R, Lingley AJ, McIsaac DI, Pysyk C, McConnell MC, Bryson GL, et al. Logging in: a comparative analysis of electronic health records versus anesthesia resident-driven logbooks. *Can J Anaesth.* 2020;**67**(10):1381-8. [PubMed ID: 32661721]. https://doi.org/10.1007/s12630-020-01761-x.