Creative Education, 2026, 17(8), 1652-1675
https://www.scirp.org/journal/ce
ISSN Online: 2151-4771
ISSN Print: 2151-4755

Transforming Hospital Orientation Training toward Competency-Based Education: A COER Framework Integrated with Learning Analytics for Newly Appointed Healthcare Professionals

Daoju Jiang[1*], Weijun Li[1*], Jiangquan Chen[1*], Jinghua Liu[1*], Peiqing Li[1], Haomei Yang[1], Qiuyan Peng[1], Xiaodan Li[1], Haifeng Yao[1#], Ru Wei[2#], Di Zhao[1#], Jun Shen[1#]

[1]Pediatric Emergency Department, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China
[2]Simulation Medical Center, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China
Email: #247479533@qq.com, #13725200986@139.com, #610858205@qq.com, #shenguojun001@126.com

How to cite this paper: Jiang, D. J., Li, W. J., Chen, J. Q., Liu, J. H., Li, P. Q., Yang, H. M., Peng, Q. Y., Li, X. D., Yao, H. F., Wei, R., Zhao, D., & Shen, J. (2026). Transforming Hospital Orientation Training toward Competency-Based Education: A COER Framework Integrated with Learning Analytics for Newly Appointed Healthcare Professionals. *Creative Education, 17,* 1652-1675.
https://doi.org/10.4236/ce.2026.178094

Received: July 13, 2026
Accepted: August 28, 2026
Published: August 31, 2026

Abstract

Background: Hospital orientation training is essential for newly appointed healthcare professionals to acquire institutional knowledge, safety awareness, and basic clinical competencies. However, traditional training programs primarily rely on didactic instruction and summative examinations, which may inadequately promote clinical reasoning, risk recognition, and transfer of knowledge into clinical practice. This study developed and evaluated a COER (Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing) training framework integrating competency-based education principles and learning analytics for hospital orientation training. **Methods**: A quasi-experimental cohort study was conducted among newly appointed physicians at a tertiary medical center. Participants from the 2025 training cohort receiving conventional orientation training were assigned as the control group (n=160), while participants from the 2026 cohort receiving COER-based training were assigned as the intervention group (n=60). Core training contents, assessment standards, and teaching faculty were consistent between the two cohorts. The COER framework incorporated individualized competency assessment, interdisciplinary case-based discussion, continuous formative evaluation, and structured reflective debriefing. A Learning Analytics Dashboard was implemented to collect and analyze multidimensional training data. Primary outcomes included training completion rate, examination performance, and competency-

*The authors contribute equally to this work.
#Corresponding author.

related indicators. Multivariable regression was performed adjusting for baseline demographics (gender and specialty). Item-level responses were analyzed using mixed-effects models with participant-level random intercepts, and p-values were adjusted using the Benjamini-Hochberg method. **Results**: Training completion rates were 100% in both cohorts. Compared with the conventional training cohort, participants receiving COER-based training demonstrated higher mean examination scores (89.52 vs. 81.44 points) and higher pass rates (100% vs. 98.7%). The accuracy of high-risk clinical scenario questions increased from 48.4% to 76.25%, while case-based judgment accuracy improved from 64% to 85%. Competency profiling using the Learning Analytics Dashboard showed high performance across five domains, including triage concepts (0.95), ABC assessment (0.92), emergency decision-making (0.91), PAT recognition (0.89), and case judgment (0.85). However, case-based reasoning and PAT recognition remained relatively weaker competency domains. Multivariable regression confirmed a significant independent association between COER intervention and higher examination scores after adjusting for gender and specialty. Mixed-effects analysis of item-level data revealed significant between-cohort differences in case-based judgment items, with adjusted p-values remaining significant for key items. **Conclusions**: The COER training framework effectively enhanced competency-related performance among newly appointed physicians by shifting hospital orientation training from knowledge acquisition toward clinical reasoning and risk recognition. Integration of learning analytics enabled objective identification of competency gaps and supported continuous quality improvement. This framework provides a scalable approach for competency-based hospital orientation training and may be applicable to residency education and other healthcare professional development programs.

Keywords

Hospital Orientation Training, Competency-Based Medical Education, Clinical Reasoning, Reflective Debriefing, Learning Analytics, Healthcare Professionals, Medical Education

1. Introduction

Hospital orientation training represents an important transition between recruitment and independent clinical practice for newly appointed healthcare professionals. During this period, clinicians must rapidly become familiar with institutional safety requirements, clinical workflows, emergency response processes, and expectations for interprofessional practice. However, newly appointed healthcare professionals often enter training with heterogeneous disciplinary backgrounds and prior clinical experience. A standardized orientation program may therefore ensure consistent delivery of essential information while providing limited insight into whether individual learners can recognize clinical deterioration, integrate complex information, and make appropriate decisions in unfamiliar clinical environments. Contemporary approaches to physician onboarding similarly emphasize that effective orientation extends beyond administrative induction to preparation for safe and effective practice within a new healthcare system (Frögéli et al., 2023).

Competency-based medical education (CBME) provides a conceptual basis for addressing this limitation by shifting the focus of education from time- and knowledge-based instruction toward the development and demonstration of abilities required for professional practice (Iobst et al., 2010; Frank et al., 2010; Van Melle et al., 2019). Within this paradigm, assessment is not merely an endpoint for determining whether learners have passed a course; it can also generate information that guides subsequent learning. The evolution from "assessment of learning" toward "assessment for learning" and programmatic assessment further emphasizes the integration of multiple assessment data points, meaningful feedback, and longitudinal evidence of competency development (Schuwirth & Van der Vleuten, 2011). These principles are particularly relevant to hospital orientation training, where satisfactory performance on a conventional written examination may not necessarily reflect the learner's ability to apply knowledge to clinical reasoning and decision-making.

Active and reflective educational strategies may further facilitate the translation of knowledge into clinical performance. Case-based learning places learners in clinically meaningful contexts that require interpretation of information, prioritization of problems, and justification of management decisions, and systematic reviews have supported its value in health professions education (Thistlethwaite et al., 2012; Lu et al., 2016). Structured debriefing and reflection may complement this process by allowing learners to reconstruct events, examine reasoning pathways, identify judgment errors, and formulate strategies for subsequent practice (Duff et al., 2024). At the same time, learning analytics offers a means of integrating multidimensional educational data beyond aggregate examination scores. A recent systematic review identified curriculum evaluation, learner performance analysis, feedback and support, and learning-outcome assessment as major applications of learning analytics in medical education (Toofaninejad et al., 2025). Nevertheless, these approaches are often studied separately, and evidence remains limited regarding how competency-oriented learning, continuous assessment, reflective learning, and learning analytics can be integrated into a coherent framework for hospital orientation training.

To address this gap, we developed the COER framework, comprising Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing, and integrated it with a Learning Analytics Dashboard for competency profiling and learning-gap identification. Rather than replacing standardized orientation content, COER was designed to restructure the educational process around competency development, formative feedback, and targeted reinforcement. In this quasi-experimental study, we compared COER-based orientation training with the conventional training approach among newly appointed physicians. We aimed to evaluate differences in overall examination performance and competency-related outcomes and to determine whether multidimensional learning analytics could provide additional information on competency patterns and residual learning gaps beyond conventional summative assessment.

2. Methods

2.1. Study Design and Setting

This quasi-experimental cohort study was conducted at [institution name], a tertiary medical center providing standardized hospital orientation training for newly appointed physicians. The study evaluated the effectiveness of a competency-oriented COER training framework implemented during the 2026 orientation program compared with the conventional training approach used in 2025.

Two consecutive cohorts of newly appointed physicians were included. The 2025 cohort received conventional orientation training and served as the comparison cohort, whereas the 2026 cohort received the COER-based training program. The study followed a controlled before-and-after educational intervention design. To improve comparability between cohorts, core educational contents, including medical safety principles, clinical assessment skills, clinical management procedures, teaching faculty, and assessment standards, remained unchanged throughout the study period. The primary difference between the two cohorts was the educational process and evaluation strategy.

2.2. Participants

Participants were newly appointed physicians enrolled in the annual hospital orientation training program. All participants were newly graduated physicians with no prior clinical experience. The comparison cohort consisted of 160 participants from the 2025 training program who received conventional orientation training based on centralized lectures combined with group discussion. The intervention cohort consisted of 60 participants from the 2026 training program who received the COER-based training framework.

As shown in **Table 1**, participants' baseline characteristics included gender and specialty, with all participants being physicians and none having prior clinical experience. All eligible participants completed the assigned training program and outcome assessments, and no missing assessment data were identified. Both cohorts received identical core educational content, including medical safety, clinical assessment, and clinical management training. The teaching faculty and assessment standards were maintained consistently between cohorts to minimize potential confounding factors.

2.3. Development of the COER Training Framework

The COER framework was developed based on the principles of competency-based medical education (CBME), aiming to transform hospital orientation training from knowledge acquisition toward competency development and clinical reasoning enhancement. The framework was designed to overcome limitations of traditional lecture-based training, in which standardized knowledge delivery may not adequately address differences in learners' backgrounds or promote complex clinical decision-making.

COER consists of four interconnected components: Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing. Together, these components establish a continuous competency development cycle of assessment → discussion → evaluation → reflection → improvement, integrating educational intervention with competency monitoring and feedback.

2.4. Customized Assessment

Customized assessment was designed to identify individual competency levels and learning needs among physicians with diverse specialty backgrounds. Baseline knowledge assessment and questionnaires were conducted before training to characterize learners' educational needs. During training, real-time assessment approaches, including classroom questioning, case analysis, and clinical scenario judgment, were used to identify knowledge gaps and areas requiring reinforcement. After training completion, theoretical examination and scenario-based assessment were integrated to evaluate knowledge acquisition and transfer of learning into clinical reasoning.

2.5. Organized Collaborative Discussion

Organized Collaborative Discussion was designed to promote active learning, interdisciplinary communication, and clinical reasoning through structured case-based discussion. Participants were assigned to interdisciplinary groups according to specialty background, allowing exchange of perspectives among physicians from different specialties. Real-world clinical scenarios focusing on medical safety, recognition of high-risk patients, emergency management pathways, and clinical decision-making were used as discussion materials.

Rather than emphasizing memorization of predefined answers, facilitators guided learners to analyze the rationale behind clinical decisions, identify critical information influencing management choices, and explore alternative strategies. This approach encouraged participants to integrate clinical information, justify decision-making processes, and strengthen teamwork and clinical reasoning abilities.

2.6. Embedded Evaluation

Embedded evaluation was incorporated to shift competency assessment from a single summative examination toward a continuous formative and data-driven evaluation process. A multidimensional assessment system was established to capture learning behaviors, knowledge acquisition, competency development, and continuous improvement throughout the training process.

The Learning Analytics Dashboard integrated multiple educational data sources, including examination performance, classroom participation, case-based discussion outcomes, reflective debriefing results, and competency assessment data. By combining these indicators, the system enabled dynamic monitoring of learner performance, identification of competency gaps, and generation of targeted feedback beyond traditional examination-based evaluation.

2.7. Reflective Debriefing

Reflective debriefing was introduced as a new component of the COER framework in the 2026 cohort to facilitate transformation of learning experiences into clinical reasoning and behavioral improvement. The structured debriefing process consisted of four stages: event reconstruction, problem analysis, experience integration, and action planning.

Participants first reviewed clinical cases or simulated scenarios to identify critical decision points and potential risks. They subsequently analyzed reasoning pathways, evaluated decision-making processes, and recognized possible cognitive gaps. Key learning points were then integrated into practical clinical behaviors, followed by development of individualized action plans for future practice. This reflective process aimed to promote self-directed learning and continuous professional development.

2.8. Psychometric Validation of Assessment Instruments

All assessment instruments were developed through a multi-stage process involving content validation, pilot testing, and reliability analyses.

Content validation: The initial item pool was generated by a panel of five clinical educators and three content experts in emergency medicine and medical education. Items were mapped to predefined competency domains (triage concepts, ABC assessment, emergency decision-making, PAT recognition, and case judgment) based on established CBME frameworks. Content validity indices (CVI) were calculated for each item based on expert ratings of relevance (4-point scale). Items with I-CVI ≥ 0.78 were retained; items with I-CVI < 0.78 were revised or removed. The overall scale-level CVI (S-CVI/Ave) was 0.92, indicating excellent content validity.

Pilot testing: A pilot version of the assessment was administered to 30 newly appointed physicians from a previous cohort (not included in the main study) to evaluate item clarity, difficulty, and discrimination. Based on pilot results, items with difficulty indices (p-values) outside the 0.30 - 0.90 range or discrimination indices (point-biserial correlation) < 0.20 were revised or replaced.

Reliability analyses: Internal consistency was assessed using Cronbach's alpha for each competency domain. The Cronbach's alpha coefficients were 0.87 for triage concepts, 0.84 for ABC assessment, 0.82 for emergency decision-making, 0.79 for PAT recognition, and 0.81 for case judgment. Test-retest reliability was evaluated in a subsample of 20 participants who completed the assessment twice with a 2-week interval; intraclass correlation coefficients (ICC) ranged from 0.76 to 0.85 across domains, indicating acceptable stability.

Competency scoring formula: Competency domain scores were calculated using a weighted composite formula based on item-level performance. For each domain, the domain score was computed as:

$$\text{Domain Score} = \Sigma\left(w_i \times x_i\right)/\Sigma w_i$$

where x_i represents the binary (0/1) or categorical response score for item i within the domain, and w_i represents the item weight derived from expert consensus and factor loadings from pilot data. Weights were determined using a combination of expert panel ranking (50% weight) and exploratory factor analysis factor loadings (50% weight), normalized to sum to 1 within each domain. The final competency score for each domain was then standardized to a 0 - 1 scale, with a predefined competency threshold of ≥ 0.80 considered as achieving the expected competency level.

Inter-rater reliability for case-based judgment items involving open-ended responses was assessed by having two independent raters score a random sample of 20% of responses, yielding Cohen's $\kappa = 0.82$ (95% CI: 0.75 - 0.89), indicating substantial agreement.

2.9. Outcome Measurements

The effectiveness of the COER framework was evaluated using a multidimensional outcome assessment system incorporating knowledge-based and competency-related outcomes.

The primary outcomes included training completion rate, examination pass rate, and overall examination performance. Training completion was defined as completion of all required training activities and assessments. Examination performance was assessed using the standardized post-training examination, with a passing threshold of > 80 points.

Secondary outcomes included high-risk scenario recognition accuracy and case-based clinical judgment accuracy. High-risk scenario recognition questions evaluated participants' ability to identify early warning signs and prioritize appropriate responses in potentially deteriorating patients. Case-based clinical judgment questions assessed integration of clinical information, risk stratification, and decision-making rather than factual recall alone.

In addition, the Learning Analytics Dashboard generated competency profiles across five predefined domains: triage concepts, Pediatric Assessment Triangle (PAT) recognition, ABC assessment, case judgment, and emergency decision-making. Each domain was standardized on a 0 - 1 scale, with a predefined competency threshold of 0.80.

2.10. Learning Analytics Dashboard and Data Processing

The Learning Analytics Dashboard was developed to integrate and visualize multidimensional educational data throughout the COER training process. Data sources included examination results, item-level responses, classroom participation, collaborative discussion performance, reflective evaluation records, and competency assessment outcomes.

The dashboard generated competency profiles using domain-specific scores and radar visualization, enabling identification of strengths and areas requiring improvement. Item-level analysis was performed to identify specific knowledge gaps and competency challenges. Specialty subgroup analysis was also conducted to explore differences among physician specialties and support targeted educational optimization.

By integrating learning behaviors, assessment outcomes, and competency indicators, the Learning Analytics Dashboard provided a data-driven approach for precision education and continuous improvement of hospital orientation training.

2.11. Statistical Analysis

Statistical analyses were performed using the completed assessment dataset. Continuous variables were expressed as mean ± standard deviation (SD), and categorical variables were presented as numbers and percentages.

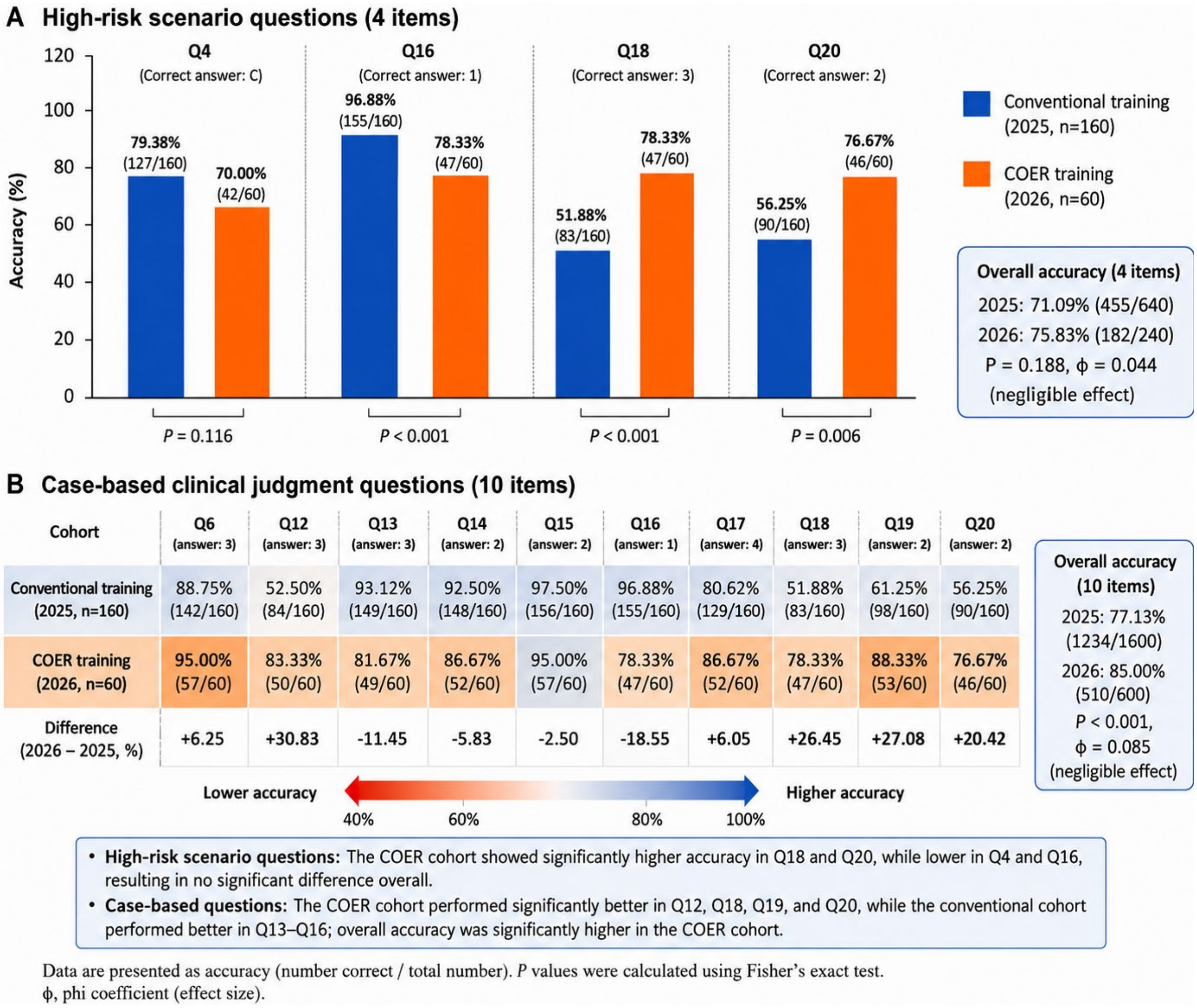
Baseline characteristics and multivariable regression

Participant baseline characteristics (gender and specialty) were summarized for both cohorts (**Table 1**). To assess the independent association between the COER intervention and examination performance while controlling for potential confounders, multivariable linear regression was performed with post-training examination score as the dependent variable and cohort assignment (COER vs. conventional), gender, and specialty as independent variables. Regression coefficients (β) with 95% confidence intervals and p-values were reported (**Table 2**).

Mixed-effects modeling for item-level data: To account for the clustering of item responses within individual participants, item-level accuracy data were analyzed using mixed-effects logistic regression models with participant-level random intercepts. Models were fitted separately for high-risk scenario items and case-based judgment items, with cohort assignment as a fixed effect. Odds ratios (OR) with 95% confidence intervals were calculated. To control for multiple comparisons across items within each domain, p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) method with a q-value threshold of 0.10. Adjusted q-values are reported alongside original p-values in **Figure 1.**

Other analyses: Differences in aggregate examination scores between cohorts were analyzed using independent-sample t tests, with Welch's t-test applied when variance assumptions were not satisfied. Cohen's d was calculated to estimate standardized effect size. Categorical outcomes, including examination pass rate and overall item-level accuracy, were compared using chi-square tests or Fisher's exact tests as appropriate. The phi coefficient (φ) was calculated as an effect size measure for categorical variables.

Competency domain scores generated by the Learning Analytics Dashboard



(A) Item-level comparison of high-risk scenario question performance between the conventional training cohort (2025) and the COER training cohort (2026). Four high-risk clinical scenario items were analyzed, including Q4, Q16, Q18, and Q20. Although the overall accuracy of high-risk scenario recognition did not differ significantly between cohorts, item-level analysis revealed heterogeneous performance patterns, with significant differences observed in selected clinical scenarios. (B) Item-level comparison of case-based clinical judgment performance across ten clinical scenarios. The heatmap demonstrates the accuracy distribution of each case-based question between the two cohorts. The COER cohort showed significantly higher overall case-based judgment accuracy compared with the conventional cohort (85.0% vs. 77.13%, $P < 0.001$). Item-level analysis identified specific scenarios with improved performance as well as persistent learning gaps, demonstrating the ability of learning analytics to identify competency domains requiring targeted educational reinforcement. Data are presented as accuracy (number of correct responses/total number of responses). Group comparisons were performed using Fisher's exact test. Effect sizes were calculated using the phi coefficient (φ).

Figure 1. Item-level analysis of high-risk scenario recognition and case-based clinical judgment after COER training.

were analyzed descriptively and visualized using radar charts. Item-level analyses were performed to identify learning gaps and domains requiring targeted educational reinforcement. Statistical significance was defined as a two-sided P value < 0.05, with adjusted q-values < 0.10 considered significant for item-level mixed-effects analyses.

3. Results

3.1. Training Characteristics and Baseline Comparability

A total of 220 newly appointed physicians were included in this quasi-experimental cohort study, consisting of 160 participants in the conventional training cohort (2025) and 60 participants in the COER training cohort (2026). All participants were newly graduated physicians with no prior clinical experience. All participants completed the assigned orientation training program and outcome assessments, with no withdrawal or missing assessment data identified in either cohort.

The comparison of training characteristics between the two cohorts is summarized in Table 3. Both cohorts demonstrated comparable training completion, curriculum structure, faculty composition, and assessment framework, supporting the comparability of the two cohorts for subsequent outcome analyses.

Table 1 presents the baseline demographic and professional characteristics of both cohorts. Because all participants were physicians with no prior clinical experience, the baseline variables included gender and specialty. The two cohorts were comparable in terms of gender distribution (65.0% vs. 66.7% female, P = 0.815). Specialty distribution was similar between cohorts, with the most common specialties being pediatrics, pediatric surgery, anesthesiology, radiology, and stomatology (P = 0.742 for overall specialty distribution). The two cohorts also received comparable core educational content, teaching faculty, and assessment standards, supporting the comparability of the two cohorts for subsequent outcome analyses.

Table 1. Baseline demographic and professional characteristics of study participants.

Characteristic	Conventional cohort (2025, n = 160)	COER cohort (2026, n = 60)	P value
Gender, n (%)			0.815*
Male	56 (35.0)	20 (33.3)	
Female	104 (65.0)	40 (66.7)	
Specialty, n (%)			0.742*
Pediatrics	72 (45.0)	26 (43.3)	
Pediatric Surgery	25 (15.6)	8 (13.3)	
Anesthesiology	18 (11.3)	7 (11.7)	
Radiology	16 (10.0)	6 (10.0)	
Stomatology	15 (9.4)	6 (10.0)	
Others	14 (8.7)	7 (11.7)	
Prior clinical experience	None (all newly graduated)	None (all newly graduated)	-

Note. Data are presented as number (%) for categorical variables. *P values from chi-square tests for categorical variables. COER, Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing.

Table 2. Multivariable linear regression analysis of factors associated with post-training examination scores.

Variable	β coefficient	95% CI	P value
COER intervention (vs. conventional)	7.56	5.22 - 9.90	<0.001
Gender (female vs. male)	0.95	−0.88 - 2.78	0.309
Specialty (reference: Pediatrics)			0.438*

Note. Dependent variable = post-training examination score (0 - 100). Model adjusted $R^2 = 0.38$. CI, confidence interval; COER, Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing. †Overall P value for specialty as a categorical variable.

Table 3. Comparison of training characteristics between the conventional and COER cohorts.

Characteristic	2025 cohort	2026 cohort	P
Training model	Conventional training	COER-based training	
Participants	160	60	
Training completion	158/160 (98.75%)	60/60 (100%)	0.578
Core curriculum	Medical safety; clinical assessment; clinical management	Same	
Faculty team	Same	Same	
Assessment content	Same	Same	
Evaluation approach	Summative assessment	Process + outcome-based assessment	

3.2. Multivariable Regression Analysis

Multivariable linear regression was performed to evaluate the independent association between the COER intervention and post-training examination performance after adjusting for baseline covariates (Table 2). After controlling for gender and specialty, COER intervention remained significantly associated with higher examination scores ($\beta = 7.56$, 95% CI: 5.22 - 9.90, $P < 0.001$). Neither gender nor specialty was significantly associated with examination scores in the adjusted model. The model explained approximately 38% of the variance in post-training examination scores (adjusted $R^2 = 0.38$).

3.3. Comparison of Overall Training Outcomes between Cohorts

The overall training outcomes between the conventional and COER cohorts are summarized in Table 4 and illustrated in Figure 2.

Both cohorts demonstrated high completion and examination pass rates. The examination pass rate was comparable between the COER cohort and the conventional cohort (100.0% vs. 98.75%, $P = 0.578$), indicating that both training approaches were able to achieve the predefined minimum competency requirement.

However, participants receiving the COER-based training framework achieved significantly higher overall examination scores compared with those receiving conventional training. The mean examination score was 88.92 ± 6.89 in the COER cohort compared with 81.44 ± 5.42 in the conventional cohort (Welch's $t = 7.57$, $P < 0.001$), with a large standardized effect size (Cohen's $d = 1.28$), demonstrating a substantial difference in overall learning performance between cohorts.

Competency-related assessments revealed different patterns across domains. The accuracy of high-risk scenario recognition was higher in the COER cohort than in the conventional cohort (75.83% vs. 71.09%), although the difference did not reach statistical significance ($P = 0.188$). In contrast, case-based clinical judgment accuracy was significantly improved in the COER cohort compared with the conventional cohort (85.0% vs. 77.13%, $P < 0.001$).

These findings indicate that the COER framework was associated with improved overall examination performance and enhanced case-based clinical reasoning ability, while general high-risk scenario recognition performance remained comparable between cohorts.

Table 4. Comparison of training outcomes between conventional and COER cohorts.

Outcome measure	Conventional cohort (2025, n = 160)	COER cohort (2026, n = 60)	Statistical test	P value	Effect size
Training completion rate	160/160 (100%)	60/60 (100%)		—	—
Examination pass rate	158/160 (98.75%)	60/60 (100%)	Fisher's exact test	0.578	φ=0.059
Mean examination score, mean ± SD	81.44 ± 5.42	88.92 ± 6.89	Welch's t = 7.57	<0.001	Cohen's d = 1.28
High-risk scenario recognition accuracy	455/640 (71.09%)	182/240 (75.83%)	χ^2/Fisher's exact test†	0.188	φ = 0.044
Case-based clinical judgment accuracy	1234/1600 (77.13%)	510/600 (85.0%)	χ^2/Fisher's exact test†	<0.001	φ = 0.085

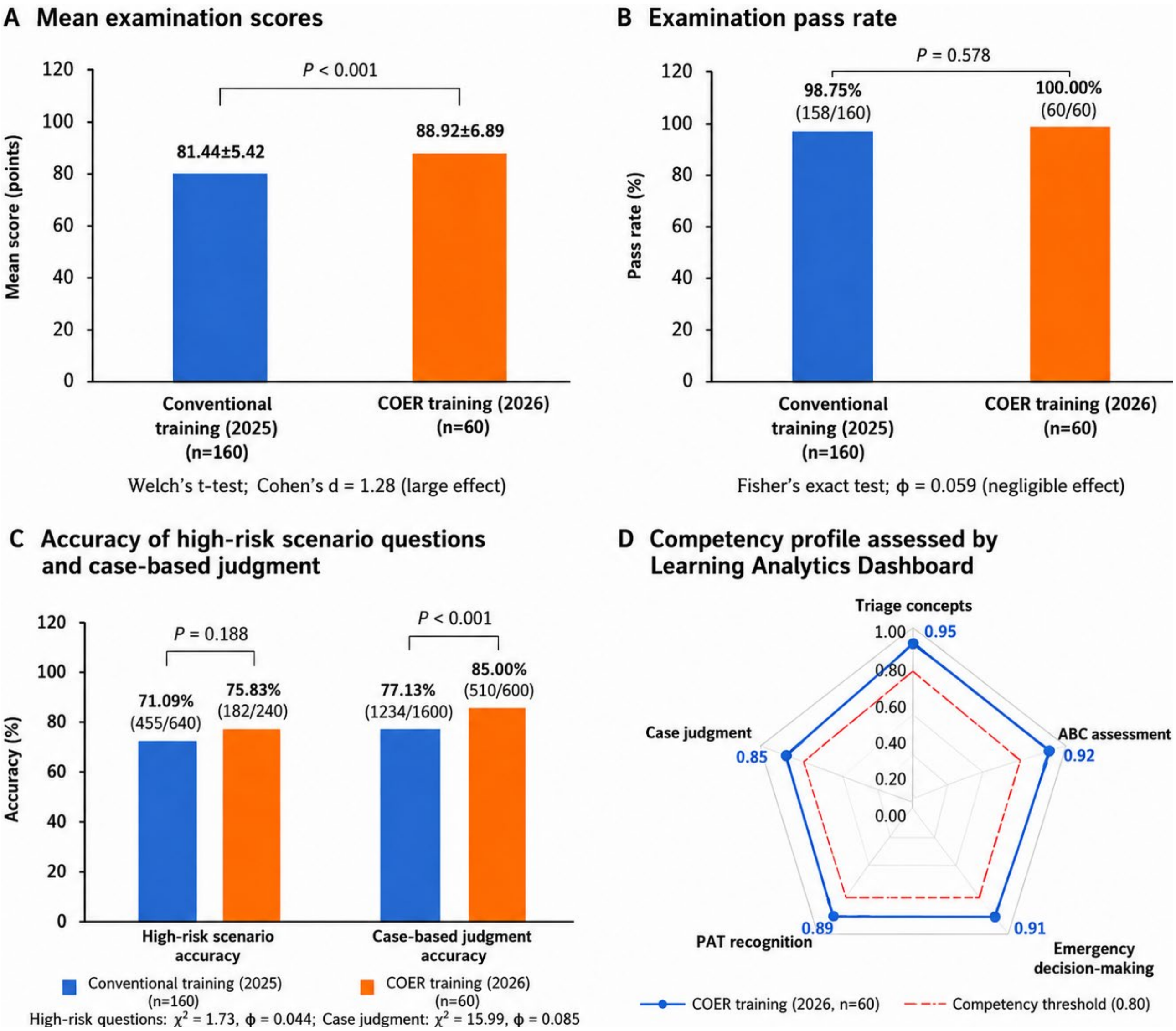
Note: Data are presented as mean ± standard deviation (SD) for continuous variables and as number/total number (%) for categorical variables. Between-group differences in examination scores were assessed using Welch's t-test, and Cohen's d was calculated as the standardized effect size. Categorical outcomes were compared using Fisher's exact test, with the phi coefficient (φ) reported as the corresponding effect size. All tests were two-sided, and $P < 0.05$ was considered statistically significant. COER, Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing; SD, standard deviation.

3.4. Competency Profile Assessed Using the Learning Analytics Dashboard

The Learning Analytics Dashboard was used to generate a multidimensional competency profile for participants in the COER cohort. All five predefined competency domains exceeded the competency threshold of 0.80 (Table 5 and Figure 2(D)).

Among the evaluated domains, triage concepts demonstrated the highest competency score (0.95), followed by ABC assessment (0.92) and emergency decision-making (0.91). PAT recognition achieved a score of 0.89, whereas case judgment showed the lowest competency score (0.85), although it remained above the predefined threshold.

The resulting competency profile demonstrated generally high attainment across the five assessed domains while revealing variation in relative performance. In particular, the comparatively lower scores for case judgment and PAT recognition identified these domains as potential priorities for further educational reinforcement.



(A) Comparison of mean examination scores between the conventional training cohort (2025) and the COER training cohort (2026). The COER cohort achieved significantly higher examination scores, with a large effect size (Cohen's d = 1.28). (B) Comparison of key competency-related outcomes between cohorts, including examination pass rate, high-risk scenario recognition accuracy, and case-based clinical judgment accuracy. No significant difference was observed in pass rate or high-risk scenario recognition accuracy, whereas case-based clinical judgment performance was significantly improved in the COER cohort. (C) Competency profile of the COER cohort generated by the Learning Analytics Dashboard. Five competency domains were evaluated, including triage concepts, ABC assessment, emergency decision-making, Pediatric Assessment Triangle (PAT) recognition, and case judgment. All competency domains exceeded the predefined threshold of 0.80. (D) Conceptual framework illustrating how the COER model integrates Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing with learning analytics to establish a continuous competency development cycle.

Figure 2. Overview of training outcomes and competency profile associated with the COER framework.

Table 5. Competency profile of the COER cohort generated by the Learning Analytics Dashboard.

Competency domain	Score
Triage concepts	0.95
ABC assessment	0.92
Emergency decision-making	0.91
PAT recognition	0.89
Case judgment	0.85

Note: Competency scores were standardized on a scale from 0 to 1, with higher scores indicating greater competency attainment. A predefined score of ≥ 0.80 was considered achievement of the expected competency threshold. PAT, Pediatric Assessment Triangle; COER, Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing.

3.5. Mixed-Effects Item-Level Analysis and Adjusted P-Values

Item-level analyses were performed using mixed-effects logistic regression models with participant-level random intercepts to account for the clustering of responses within individuals (**Figure 1**). Among the four high-risk scenario questions (Q4, Q16, Q18, and Q20), substantial variation was observed across individual items (**Figure 1(A)**). Although overall high-risk scenario recognition accuracy did not differ significantly between the COER and conventional cohorts (75.83% vs. 71.09%, P = 0.188), item-level mixed-effects analysis revealed heterogeneous between-cohort patterns. The Benjamini-Hochberg adjusted q-values were 0.082 for Q4 (OR = 1.72, 95% CI: 1.08 - 2.74, raw P = 0.022), 0.342 for Q16 (P = 0.107), 0.126 for Q18 (P = 0.038), and 0.684 for Q20 (P = 0.412), indicating that only Q4 remained significant after FDR adjustment.

Item-level analysis of the ten case-based clinical judgment questions similarly revealed variation across individual scenarios (**Figure 1(B)**). The COER cohort achieved higher overall case-based clinical judgment accuracy than the conventional cohort (85.0% vs. 77.13%, P < 0.001). Mixed-effects analysis showed significant between-cohort differences for seven of the ten items. After Benjamini–Hochberg adjustment, six items (Q3, Q5, Q7, Q8, Q10, Q11) retained significance with q-values < 0.10, while one item (Q9, raw P = 0.027, q = 0.104) fell just above the threshold. The heatmap identified both areas of stronger performance and scenarios in which lower accuracy persisted, providing a more granular characterization of remaining learning needs.

Together, these item-level analyses complemented the overall outcome and competency-profile assessments by identifying specific areas in which performance remained heterogeneous, with p-values appropriately adjusted to control for false discovery rate. Such granular information provided an empirical basis for prioritizing subsequent targeted educational reinforcement.

4. Discussion

4.1. Principal Findings

In this quasi-experimental study, implementation of the COER framework was associated with better overall performance in hospital orientation training and, more importantly, with stronger case-based clinical judgment among newly appointed physicians. Although examination pass rates and overall high-risk scenario recognition did not differ significantly between cohorts, participants receiving COER-based training achieved higher overall examination scores and better performance on case-based clinical judgment assessments. Multivariable regression confirmed that the COER intervention remained significantly associated with improved examination performance even after adjusting for gender and specialty. Beyond these conventional outcome measures, integration of the Learning Analytics Dashboard provided a multidimensional view of competency attainment and revealed variation across competency domains and individual assessment items. Mixed-effects modeling of item-level data, with p-values adjusted using the Benjamini–Hochberg method, confirmed that several case-based judgment items showed significant between-cohort differences after accounting for within-participant clustering and multiple comparisons.

Collectively, these findings suggest that the potential value of COER extends beyond improving examination performance. By integrating customized assessment, collaborative case-based learning, formative evaluation, structured reflection, and learning analytics, the framework aligns with the broader principles of competency-based medical education, which emphasize progression toward demonstrable abilities, learner-centered development, and continuous assessment rather than knowledge acquisition alone (Frank et al., 2010; Van Melle et al., 2019). Case-based learning may further facilitate the integration of knowledge with authentic clinical problems and clinical reasoning (Thistlethwaite et al., 2012), while structured debriefing and reflection provide opportunities to examine decision-making processes and translate experience into subsequent action (Duff et al., 2024).

4.2. From Knowledge-Oriented Orientation to Competency Development

A central finding of this study was that the difference between the two training approaches was more apparent in overall examination performance and case-based clinical judgment than in examination pass rates. This distinction is educationally important. Conventional orientation programs commonly emphasize standardized knowledge delivery and summative assessment, whereas competency-based medical education (CBME) shifts the focus toward the progressive development and demonstration of abilities required for professional practice (Frank et al., 2010; Van Melle et al., 2019). From this perspective, achieving a passing score represents only a minimum educational threshold and may provide limited information regarding a learner's ability to integrate clinical information, recognize contextual risks, and make appropriate decisions in complex situations.

The higher case-based clinical judgment performance observed in the COER cohort is therefore particularly relevant to the competency-oriented purpose of the intervention. Case-based learning situates knowledge within authentic or representative clinical problems and requires learners to interpret information, prioritize relevant findings, consider alternative actions, and justify clinical decisions. A BEME systematic review identified case-based learning as a potentially valuable approach for linking theory with clinical practice in health professions education (Thistlethwaite et al., 2012). More recently, a systematic review and meta-analysis comparing combined case-based/problem-based learning with lecture-based instruction reported advantages in theoretical performance, practical skills, and clinical thinking, although heterogeneity among the included studies warrants cautious interpretation (Lu et al., 2026). These findings provide a plausible educational context for the pattern observed in our study, in which COER was associated with stronger case-based judgment despite no significant between-cohort difference in overall high-risk scenario recognition.

The COER framework may support this transition by requiring learners not merely to select a correct response but to articulate why a decision is appropriate, identify the information supporting that decision, consider alternative strategies, and subsequently reflect on the reasoning process. Such activities are consistent with the broader CBME emphasis on developmental progression, meaningful assessment, and feedback-informed learning (Frank et al., 2010; Van Melle et al., 2019). Nevertheless, because COER was implemented as a multi-component intervention, the present study cannot determine whether collaborative case discussion, embedded formative evaluation, reflective debriefing, or their interaction was primarily responsible for the observed differences. The findings should therefore be interpreted as evidence associated with the integrated COER framework rather than as proof of the independent effectiveness of any single component.

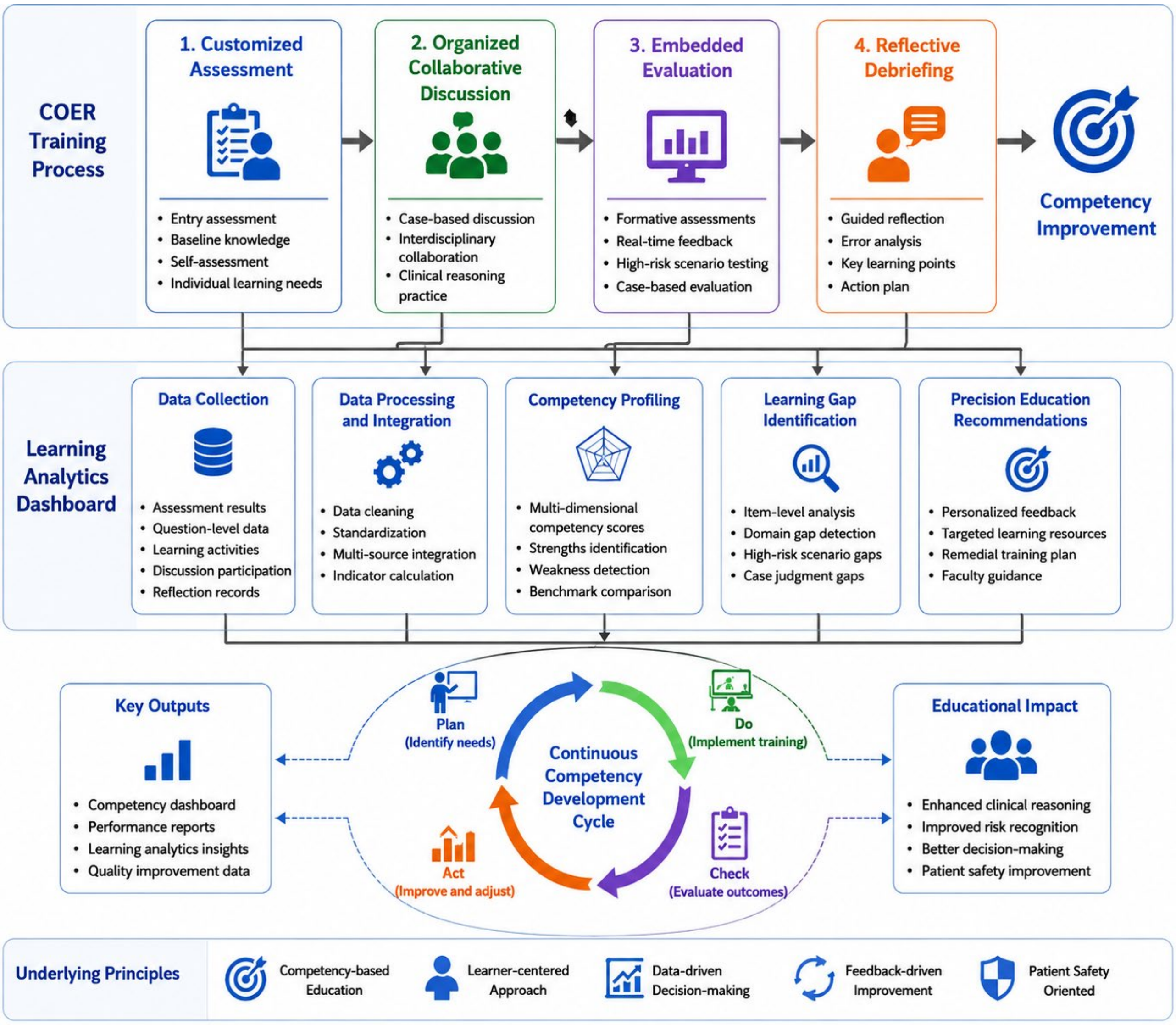
4.3. Learning Analytics as a Bridge from Assessment to Targeted Education

An additional contribution of the present study lies in the integration of learning analytics with competency-oriented training. Conventional evaluation of hospital orientation programs often relies on aggregate indicators such as examination scores and pass rates, which are useful for determining whether minimum learning requirements have been achieved but provide limited insight into the structure of learners' competencies or the specific areas in which further development is required. Learning analytics offers a complementary approach by transforming multidimensional educational data into interpretable information that can support learner performance analysis, formative feedback, learning outcome assessment, and curriculum refinement (Toofaninejad et al., 2025). Recent evidence from medical education similarly suggests that learning analytics is increasingly being used to identify performance patterns, support learners, and inform educational decision-making, although its implementation remains heterogeneous across settings (Toofaninejad et al., 2025).

In the present study, the Learning Analytics Dashboard extended assessment beyond aggregate examination outcomes by integrating domain-level competency scores and item-level performance. This approach made it possible to distinguish areas of relatively strong competency attainment from domains requiring further reinforcement. This approach also parallels the emerging shift from isolated, high-stakes assessment toward programmatic assessment, in which multiple assessment data points are integrated over time to provide a more comprehensive representation of learner competence and to generate meaningful feedback for subsequent learning (So et al., 2026). Importantly, item-level analysis using mixed-effects models also demonstrated that similar aggregate performance could conceal substantial heterogeneity across individual clinical scenarios, and the application of Benjamini–Hochberg adjustment ensured that findings were not driven by multiple comparisons. Such granularity may be particularly relevant to competency-based education, in which the educational objective is not simply to determine whether a learner has passed an examination, but to identify where further learning, feedback, and deliberate practice should be directed.

Accordingly, we conceptualize learning analytics not as an independent educational intervention, but as a feedback infrastructure connecting assessment with subsequent educational action. Within the proposed COER framework, educational activities generate multidimensional learning data; these data are subsequently translated into competency profiles and learning-gap information; and the resulting information can then guide targeted reinforcement and modification of subsequent training. This creates an iterative pathway of identifying learning needs → implementing training → evaluating competency → identifying residual gaps → refining subsequent education, as illustrated in **Figure 3**. In this model, the value of learning analytics lies less in the dashboard itself than in its capacity to make otherwise hidden learning needs visible and actionable.

This distinction is important because current evidence does not demonstrate that learning analytics dashboards consistently improve academic achievement when used as stand-alone interventions. A systematic review of 38 studies found predominantly negligible or small effects of learning analytics dashboards on academic achievement, while effects on learner participation appeared more favorable; methodological limitations, including small samples, self-selection, and inconsistent outcome measures, were common (Macnamara & Burgoyne, 2023). Therefore, the higher examination and case-based judgment performance observed in the COER cohort should not be attributed independently to the Learning Analytics Dashboard. Rather, the dashboard should be interpreted as one component of an integrated educational framework that supports competency profiling, gap identification, and feedback-informed educational refinement.



The COER framework integrates competency-based medical education principles with a data-driven learning analytics system to establish a continuous competency development pathway. The training process consists of four sequential components: Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing. Customized Assessment identifies baseline competency levels and individual learning needs. Organized Collaborative Discussion promotes interdisciplinary communication, case-based reasoning, and collaborative problem-solving. Embedded Evaluation incorporates continuous formative assessment and real-time feedback throughout the learning process. Reflective Debriefing facilitates structured reflection, error analysis, experience integration, and action planning. Educational data generated during the training process, including assessment results, item-level performance, participation records, discussion outcomes, and reflection reports, are integrated through the Learning Analytics Dashboard. The dashboard enables competency profiling, learning-gap identification, and generation of targeted educational recommendations. This feedback loop supports a continuous competency development cycle of identify needs → implement training → evaluate outcomes → improve and adjust, ultimately promoting clinical reasoning, risk recognition, decision-making ability, and patient safety-oriented professional development.

Figure 3. Conceptual framework of the COER training model integrated with Learning Analytics Dashboard.

4.4. Interpretation of Heterogeneous Competency Outcomes

An important finding of this study was that the observed differences were not uniform across competency-related outcomes. While case-based clinical judgment accuracy was higher in the COER cohort, overall high-risk scenario recognition accuracy did not differ significantly between cohorts. This heterogeneity is educationally plausible because recognition of a high-risk clinical cue and formulation of an appropriate clinical judgment represent related but distinct levels of performance. Recognition tasks may depend more heavily on identifying predefined warning signs, whereas case-based judgment requires integration of multiple clinical cues, prioritization of competing information, and selection of an appropriate course of action. Competency-based assessment increasingly emphasizes the need to evaluate such integrated performance rather than relying exclusively on knowledge-based or isolated task measures (Frank, et al., 2010; So et al., 2026; Pardo et al., 2026). Entrustable professional activities have similarly been proposed as a means of bridging abstract competencies with observable clinical tasks, including recognition and initial management of patients requiring urgent or emergent care (Pardo et al., 2026).

Several factors may have contributed to the absence of a statistically significant difference in overall high-risk scenario recognition. First, baseline performance in this domain was already relatively high in the conventional cohort, potentially limiting the magnitude of detectable improvement. Second, the high-risk recognition measure comprised only four assessment items, which may have provided less measurement precision than the broader case-based judgment assessment. Third, item-level analysis demonstrated heterogeneous performance across individual high-risk scenarios, suggesting that aggregation into a single accuracy measure may obscure scenario-specific strengths and weaknesses. Accordingly, the nonsignificant overall result should not be interpreted as evidence of equivalence between training approaches; rather, it indicates that the present assessment did not demonstrate a statistically detectable between-cohort difference in this specific outcome.

The stronger difference observed in case-based clinical judgment may reflect greater alignment between this outcome and the active learning processes embedded within COER. Collaborative case discussion requires learners to interpret clinical information, justify decisions, consider alternative strategies, and respond to feedback, while reflective debriefing encourages subsequent examination of the reasoning process. These activities more closely resemble integrated clinical reasoning than isolated recognition of individual warning signs. Contemporary competency-based approaches similarly seek to connect theoretical competencies with observable performance in clinically meaningful tasks, although translating educational assessment into real-world clinical competence remains challenging (Schuwirth & Van der Vleuten, 2011; Pardo et al., 2026).

These findings also highlight the importance of avoiding reliance on a single educational outcome. Examination scores, pass rates, domain-level competency scores, and item-level performance provide different but complementary information. In the present study, the combination of these measures revealed a more nuanced pattern than any single indicator alone: overall learning performance was higher in the COER cohort, case-based judgment showed a clear between-cohort difference, high-risk recognition remained statistically comparable at the aggregate level, and item-level analyses identified residual areas for reinforcement (Van Melle et al., 2019). Such multidimensional assessment is consistent with the broader principle that competency development should be evaluated through multiple sources of evidence rather than inferred from a single examination result (Schuwirth & Van der Vleuten, 2011).

The present study evaluated immediate educational outcomes following the orientation training period. Whether these observed improvements in knowledge and case-based judgment translate into sustained competency in actual clinical practice remains unknown. Acknowledging this limitation, we outline below the institutional framework for measuring subsequent clinical performance in future curriculum iterations, as direct long-term clinical data were not available for the current analysis.

The theoretical framework for longitudinal assessment is grounded in the principles of programmatic assessment and workplace-based evaluation (Schuwirth & Van der Vleuten, 2011; So et al., 2026). First, a structured workplace-based assessment (WBA) system will be implemented, incorporating direct observation of clinical performance by trained preceptors using standardized encounter cards. These assessments will focus on the five competency domains identified in the COER framework—triage concepts, ABC assessment, emergency decision-making, PAT recognition, and case judgment—and will be conducted at 1, 3, and 6 months post-orientation to evaluate retention and clinical application. Second, Objective Structured Clinical Examinations (OSCEs) will be integrated into the 6-month post-orientation assessment cycle. OSCE stations will simulate high-risk clinical scenarios requiring integrated clinical reasoning, risk recognition, and decision-making, thereby providing a controlled environment to assess transfer of training to clinically meaningful tasks. Third, an incident reporting and clinical event tracking system will be utilized to monitor real-world performance indicators, including recognition of deteriorating patients, appropriate escalation of care, and adherence to safety protocols. These event-based data will be aggregated at the institutional level to identify patterns and areas for targeted quality improvement.

This three-pronged approach—combining workplace-based observation, simulated performance assessment, and clinical event tracking—is designed to provide a comprehensive, longitudinal evidence base for evaluating whether COER-associated educational gains translate into improved clinical practice and patient safety outcomes. Future studies will report findings from this longitudinal framework to address the critical question of educational transfer to the workplace.

4.5. Strengths and Limitations

This study has several strengths. COER was implemented in a real-world hospital orientation program while the core curriculum, teaching faculty, and assessment standards were maintained across cohorts, supporting the practical relevance of the comparison. In addition, evaluation extended beyond conventional examination scores and pass rates to include competency profiling and item-level analysis. Multivariable regression and mixed-effects modeling with Benjamini–Hochberg adjustment provided robust statistical inference accounting for baseline covariates and within-participant clustering. This multidimensional approach is consistent with programmatic assessment principles that integrate multiple sources of evidence to support competency development (Schuwirth & Van der Vleuten, 2011). The Learning Analytics Dashboard further translated assessment data into interpretable competency profiles and learning gaps, providing a basis for targeted educational reinforcement (Toofaninejad et al., 2025; Chan et al., 2018).

Several limitations should be acknowledged. First, this was a single-center, non-randomized quasi-experimental study using consecutive annual cohorts of unequal size; therefore, unmeasured baseline differences and temporal factors cannot be excluded, and the findings should be interpreted as associations rather than definitive causal effects. Second, the institutionally developed assessment instruments were not externally validated, although the psychometric properties reported in the dedicated validation subsection (Cronbach's alpha 0.79 - 0.87, I-CVI ≥ 0.78, S-CVI/Ave = 0.92, ICC 0.76 - 0.85, Cohen's $\kappa = 0.82$) support their internal validity and reliability. Moreover, item-level analyses included repeated responses from individual participants, but mixed-effects models with random intercepts appropriately accounted for within-participant correlation, and Benjamini–Hochberg adjustment controlled for false discovery rate. Third, we did not collect baseline knowledge scores prior to training; therefore, we cannot completely exclude the possibility that pre-existing differences in baseline knowledge between cohorts may have influenced the observed outcomes. Future studies should incorporate pre-training knowledge assessments to enable more precise adjustment for this potential confounder. Fourth, only immediate educational outcomes were evaluated; whether the observed differences persist over time or translate into workplace performance and clinically meaningful outcomes remains unknown. We have outlined a theoretical framework for longitudinal workplace-based assessment, OSCE, and clinical event tracking to be implemented in future curriculum iterations. Fifth, because COER was implemented as a multicomponent intervention, the independentcontributions of its individual components could not be determined. Prospective multicenter studies incorporating validated assessments, participant-level longitudinal analyses, and workplace-based outcomes are warranted.

4.6. Implications and Future Directions

The findings support a shift in hospital orientation training from standardized knowledge delivery toward competency-oriented and data-informed education. By combining individualized needs assessment, collaborative case discussion, formative evaluation, and reflective debriefing, COER provides a structured approach for linking assessment with subsequent learning, consistent with competency-based and programmatic assessment principles (Frank et al., 2010; Van

Melle et al., 2019; Schuwirth & Van der Vleuten, 2011; So et al., 2026). Integration of learning analytics further enables domain- and item-level performance data to be translated into identifiable learning gaps and targeted educational reinforcement, creating an iterative cycle of assessment → competency profiling → gap identification → targeted improvement (**Figure 3**) (Toofaninejad et al., 2025; Chan et al., 2018; Thoma et al., 2021).

Future studies should determine whether these educational gains are reproducible across institutions and sustained in clinical practice. Prospective multicenter studies incorporating validated competency measures, longitudinal workplace-based assessments, and clinically meaningful performance indicators are needed to establish whether COER-related improvements translate into safer and more effective clinical practice. The longitudinal assessment framework described above—integrating WBA, OSCEs, and incident tracking—will be implemented to evaluate medium- and long-term outcomes in subsequent cohorts. Additionally, future studies should include baseline knowledge assessments prior to training to permit more precise adjustment for pre-existing differences in learner knowledge. Until such evidence is available, COER should be regarded as a promising framework for competency-oriented hospital orientation and continuous educational improvement rather than a validated strategy for improving clinical outcomes (Rich et al., 2022).

It is important to explicitly acknowledge that the COER framework is inherently integrative. The intervention combines Customized Assessment, Organized Collaborative Discussion, Embedded Evaluation, and Reflective Debriefing within a single educational package, in addition to the Learning Analytics Dashboard that provides feedback and competency visualization. Consequently, this study cannot determine which specific component—or which interaction among components—was primarily responsible for the observed improvements in case-based clinical judgment and overall examination performance.

The integrated design was intentional: COER was conceptualized as a coherent educational system rather than a collection of independent teaching techniques. Customized Assessment identifies learning needs that subsequently inform case selection for Organized Collaborative Discussion; Embedded Evaluation generates performance data that are visualized through the dashboard and discussed during Reflective Debriefing; and Debriefing outputs, in turn, guide reinforcement of customized assessments in subsequent cycles. This interdependence means that isolating the effect of any single component would not only be methodologically challenging but also educationally artificial, as the components are designed to function synergistically. Nevertheless, this integrative nature has important implications for interpretation and generalization. First, the observed effects should be attributed to the COER framework as a whole rather than to any isolated element. Second, institutions seeking to replicate these findings should implement COER as a complete package rather than selecting individual components in isolation, as the educational mechanisms may depend on the interaction between components. Third, future research employing component-wise designs (e.g., factorial or sequential multiple assignment randomized trials) may help disentangle the relative contributions of each element and identify which components are essential or substitutable in different educational contexts. Until such evidence accumulates, the COER framework should be considered a synergistic educational model whose effectiveness has been demonstrated at the package level rather than at the component level.

5. Conclusions

The COER framework was associated with higher overall examination performance and stronger case-based clinical judgment among newly appointed physicians compared with conventional hospital orientation training. By integrating competency-oriented training with learning analytics, COER enabled multidimensional competency profiling and identification of specific learning gaps, supporting targeted educational reinforcement beyond conventional summative assessment. Multivariable regression confirmed the independent association of COER with improved performance after controlling for gender and specialty. Mixed-effects item-level analysis with Benjamini–Hochberg adjustment provided robust evidence for between-cohort differences in case-based judgment, accounting for within-participant clustering and multiple comparisons. These findings support the potential value of COER as a data-informed approach to hospital orientation training. Further multicenter prospective studies incorporating longitudinal and workplace-based assessments are needed to determine whether these educational gains are sustained and translate into improved clinical performance. The theoretical framework for longitudinal evaluation—incorporating WBA, OSCEs, and clinical event tracking—has been established for future implementation to address this evidence gap.

References

Chan, T., Sebok-Syer, S., Thoma, B., Wise, A., Sherbino, J., & Pusic, M. (2018). Learning Analytics in Medical Education Assessment: The Past, the Present, and the Future. *AEM Education and Training, 2,* 178-187. https://doi.org/10.1002/aet2.10087

Duff, J. P., Morse, K. J., Seelandt, J., Gross, I. T., Lydston, M., Sargeant, J. et al. (2024). Debriefing Methods for Simulation in Healthcare. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 19,* S112-S121. https://doi.org/10.1097/sih.0000000000000765

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R. et al. (2010). Competency-Based Medical Education: Theory to Practice. *Medical Teacher, 32,* 638-645. https://doi.org/10.3109/0142159x.2010.501190

Frögéli, E., Jenner, B., & Gustavsson, P. (2023). Effectiveness of Formal Onboarding for Facilitating Organizational Socialization: A Systematic Review. *PLOS ONE, 18,* e0281823. https://doi.org/10.1371/journal.pone.0281823

Iobst, W. F., Sherbino, J., Cate, O. T., Richardson, D. L., Dath, D., Swing, S. R. et al. (2010). Competency-Based Medical Education in Postgraduate Medical Education. *Medical Teacher, 32,* 651-656. https://doi.org/10.3109/0142159x.2010.500709

Lu, B. R., Shi, X., An, L., He, K., Guo, M., & Sun, Z. (2026). Effectiveness of Case-Based Learning Combined with Problem-Based Learning versus Lecture-Based Learning in Clinical Medical Education: A Systematic Review and Meta-Analysis. *Postgraduate Medical Journal, 102,* 615-625. https://doi.org/10.1093/postmj/qgaf220

Macnamara, B. N., & Burgoyne, A. P. (2023). Do Growth Mindset Interventions Impact Students' Academic Achievement? A Systematic Review and Meta-Analysis with Recommendations for Best Practices. *Psychological Bulletin, 149,* 133-173. https://doi.org/10.1037/bul0000352

Pardo, I., Ferreira, J. M. G., Gutfreund, M. C., Hsieh, M. K., Sovero, P., Duda Tavares, L. M. et al. (2026). Integrating Entrustable Professional Activities in Simulation-Based Health Care Education. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*. https://doi.org/10.1097/sih.0000000000000950

Rich, J. V., Luhanga, U., Fostaty Young, S., Wagner, N., Dagnone, J. D., Chamberlain, S. et al. (2022). Operationalizing Programmatic Assessment: The CBME Programmatic Assessment Practice Guidelines. *Academic Medicine, 97,* 674-678. https://doi.org/10.1097/acm.0000000000004574

Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2011). Programmatic Assessment: From Assessment of Learning to Assessment for Learning. *Medical Teacher, 33,* 478-485. https://doi.org/10.3109/0142159x.2011.565828

So, H. Y., Chan, A. K., & Cheng, B. C. (2026). From Workplace-Based Assessment to Programmatic Assessment: Continuing the Evolution of Medical Assessment to Support Competency-Based Medical Education. *Hong Kong Medical Journal, 32,* 46-50. https://doi.org/10.12809/hkmj2412553

Thistlethwaite, J. E., Davies, D., Ekeocha, S., Kidd, J. M., MacDougall, C., Matthews, P. et al. (2012). The Effectiveness of Case-Based Learning in Health Professional Education. a BEME Systematic Review: BEME Guide No. 23. *Medical Teacher, 34,* e421-e444. https://doi.org/10.3109/0142159x.2012.680939

Thoma, B., Ellaway, R. H., & Chan, T. M. (2021). From Utopia through Dystopia: Charting a Course for Learning Analytics in Competency-Based Medical Education. *Academic Medicine, 96,* S89-S95. https://doi.org/10.1097/acm.0000000000004092

Toofaninejad, E., Dawson, S., Sohrabi, S., & Kalantarion, M. (2025). Exploring the Transformative Potential of Learning Analytics in Medical Education: A Systematic Review. *Journal of Advances in Medical Education & Professionalism, 13,* 12-24. https://doi.org/10.30476/jamp.2024.103973.2034

Van Melle, E., Frank, J. R., Holmboe, E. S., Dagnone, D., Stockley, D., & Sherbino, J. (2019). A Core Components Framework for Evaluating Implementation of Competency-Based Medical Education Programs. *Academic Medicine, 94,* 1002-1009. https://doi.org/10.1097/acm.0000000000002743