

Article

Effect of Optimized Clinical Teaching Mode for Orthopedic Standardized Residency Training on Pass Rates of Two Key Examinations in a Municipal Tertiary Class-A Hospital

Gujun Xu¹, Zhengjun Luo¹, Mingkai Wang¹ and Yang Zhang^{1,*}

¹ Huludao Central Hospital, Huludao, China

* Correspondence: Yang Zhang, Huludao Central Hospital, Huludao, China

Abstract: Objective: To evaluate the impact of an optimized clinical teaching model for orthopedic standardized residency training (SRT) on trainees' first-attempt pass rate of the National Medical Licensing Examination (NMLE) and their performance in the final SRT assessment, and to provide empirical evidence for teaching reform of orthopedic residency training in municipal tertiary Class-A hospitals in China. Methods : We conducted a retrospective historical cohort study. A total of 42 SRT trainees admitted to the Department of Orthopedics at a municipal tertiary Class-A hospital between January 2017 and December 2019 were allocated to the traditional teaching group, which received routine clinical preceptorship. Another 45 trainees admitted between January 2020 and December 2022 were assigned to the optimized teaching group, which received a comprehensive teaching model integrating case-based learning (CBL), simulation skills training, formative assessment, and a mentor accountability system. Outcome measures included the first-attempt NMLE pass rate, final SRT assessment scores (theoretical, practical, and total scores), teaching satisfaction, and clinical competence evaluation. Categorical variables were compared using the chi-square test, continuous variables were compared using the independent samples t-test, and multivariate logistic regression was performed to identify factors associated with examination pass rates. Results : The first-attempt NMLE pass rate was significantly higher in the optimized teaching group (88.89%, 40/45) than in the traditional teaching group (69.05%, 29/42) ($\chi^2=5.87$, $P=0.015$). The optimized group also achieved significantly higher final SRT assessment scores, with theoretical (82.35 ± 6.21 vs. 73.48 ± 7.12), practical (85.62 ± 5.38 vs. 76.25 ± 6.89), and total scores (83.98 ± 5.76 vs. 74.86 ± 6.53) all exceeding those of the traditional group ($t=6.32$, 7.05 , and 7.58 , respectively; all $P<0.001$). Teaching satisfaction (95.56% vs. 78.57%) and the excellent rate of clinical competence evaluation (82.22% vs. 52.38%) were also significantly higher in the optimized group (both $P<0.05$). Multivariate logistic regression revealed that the optimized teaching model was an independent protective factor against NMLE failure, significantly reducing the risk of examination failure (OR=0.21, 95% CI: 0.07–0.64, $P=0.006$). Conclusion: The optimized clinical teaching model for orthopedic SRT significantly improves trainees' first-attempt NMLE pass rate and final SRT assessment scores, while also enhancing teaching satisfaction and comprehensive clinical competence. This model offers a replicable, practical approach for improving the quality of orthopedic residency training in municipal tertiary Class-A hospitals.

Received: 20 August 2026

Revised: 23 August 2026

Accepted: 04 September 2026

Published: 13 October 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: orthopedics; standardized residency training; clinical teaching mode; National Medical Licensing Examination; final assessment of standardized residency training

1. Introduction

Since the full implementation of the standardized residency training (SRT) system in China in 2014, it has served as a critical bridge connecting undergraduate medical education with clinical practice and a statutory core component of cultivating qualified clinicians [1,2]. In 2020, the General Office of the State Council of China issued the Guiding Opinions on Accelerating the Innovation and Development of Medical Education, which

explicitly called for improving the SRT system, establishing the pass rates of the National Medical Licensing Examination (NMLE) and the final SRT assessment (hereafter referred to as the "two key examinations") as core indicators for evaluating training base quality, and incorporating these metrics into the performance appraisal system for tertiary public hospitals [3].

As a core surgical specialty, orthopedics encompasses multiple subspecialties including trauma, joint, spine, and hand and foot surgery. It is characterized by a broad disease spectrum, stringent requirements for surgical proficiency, and frequent emergency scenarios, all of which place exceptionally high demands on trainees' theoretical knowledge framework, clinical technical skills, diagnostic and therapeutic decision-making capacity, and doctor-patient communication competence [4]. Clinical teaching during residency training constitutes the core setting for orthopedic talent development, and its quality directly determines whether trainees can acquire the position-required core competencies, with downstream effects on their examination performance and long-term career progression.

As regional medical centers, municipal tertiary Class-A hospitals serve as critical hubs for cultivating grassroots orthopedic clinicians, undertaking the vast majority of local orthopedic SRT workload. However, several prominent challenges persist in orthopedic SRT clinical teaching in these hospitals, directly compromising training quality [5-7]. First, teaching lacks systematic organization. Traditional preceptorship is dominated by experience-based, highly fragmented instruction, with content largely scattered across daily clinical cases and lacking a holistic curriculum design aligned with the NMLE outline and final SRT assessment standards, resulting in a disconnect between teaching and assessment. Second, practical teaching lacks sufficient depth. Trainees mostly participate in clinical work as assistants and clerks, with limited opportunities to independently perform core procedural skills such as debridement, plaster immobilization, and joint aspiration. Inadequate training volume makes it difficult to ensure procedural standardization and proficiency. Third, the evaluation and feedback mechanism is suboptimal. Assessment prioritizes summative evaluations such as the final examination, lacking whole-process formative assessment and timely feedback. Consequently, trainees' knowledge gaps and skill deficiencies cannot be identified and remediated promptly, leading to accumulated learning deficits. Finally, teaching is homogeneous and lacks stratification. A uniform preceptorship scheme is applied to trainees with diverse educational backgrounds and foundational knowledge, without differentiated instruction, resulting in marked disparities in learning outcomes.

These challenges directly contribute to low first-attempt NMLE pass rates and substantial variability in final assessment scores among orthopedic SRT trainees in municipal tertiary Class-A hospitals [6,7], which not only undermines training stability but also constrains the development of local orthopedic talent pipelines. The NMLE serves as the core threshold for medical practice entry, while the final SRT assessment constitutes the core criterion for post-entry competency. The two examinations are highly correlated and share aligned competency requirements [8]. Currently, domestic studies on orthopedic SRT teaching predominantly focus on provincial-level hospitals or single teaching methodologies. There remains a paucity of systematic teaching optimization studies that address the specialty-specific characteristics of municipal tertiary Class-A hospitals and use the pass rates of the "two key examinations" as core outcome indicators, highlighting an urgent need for reliable evidence from well-designed cohort studies.

2. Materials and Methods

2.1. General Information

A total of 87 SRT trainees admitted to the orthopedic SRT base of a municipal tertiary Class-A hospital between January 2017 and December 2022 were enrolled in this study. This retrospective historical cohort study was approved by the Hospital Ethics Committee (Ethics Approval No.: KY2026-39). All research data were anonymized, and the

requirement for written informed consent was waived due to the retrospective nature of the study.

Inclusion criteria:

1. Trainees with full-time bachelor's degree or above, who completed the full cycle of training at the orthopedic SRT base of our hospital;
2. With complete training cycle, no withdrawal, transfer to other bases, or delayed graduation;
3. No cheating or violation records in the National Medical Licensing Examination or the final SRT assessment;
4. Completed all teaching sessions and assessments in this study.

Exclusion criteria:

1. Trainees who terminated training early due to illness or personal reasons;
2. Trainees who did not take the National Medical Licensing Examination or the final SRT assessment as required;
3. Trainees with incomplete research data.

Based on admission time and the implementation timeline of the department's teaching model, trainees were divided into two groups: 42 trainees admitted between January 2017 and December 2019 were assigned to the traditional teaching group, which received routine clinical preceptorship for orthopedic SRT; 45 trainees admitted between January 2020 and December 2022 were assigned to the optimized teaching group, which received the optimized clinical teaching model. There were no statistically significant differences in baseline characteristics including gender, age, education level, and baseline theoretical score at admission between the two groups (all $P > 0.05$), confirming group comparability (Table 1).

Table 1. Comparison of Baseline Data between the Two Groups of SRT Trainees

Indicator	Traditional teaching group (n=42)	Optimized teaching group (n=45)	t/ χ^2 value	P value
Gender (male/female, n)	30/12	33/12	0.06	0.806
Age (years, $\bar{x} \pm s$)	24.5 $\pm$ 1.2	24.3 $\pm$ 1.3	0.73	0.467
Education level (bachelor's/master's, n)	28/14	31/14	0.08	0.777
Baseline theoretical score at admission (points, $\bar{x} \pm s$)	70.2 $\pm$ 5.8	71.5 $\pm$ 6.1	0.98	0.329

2.2. Research Methods

The trainees were divided into the traditional teaching group (admitted from January 2017 to December 2019) and the optimized teaching group (admitted from January 2020 to December 2022) according to admission time.

The traditional teaching group adopted the routine clinical preceptorship model for orthopedic SRT. Theoretical instruction was based on unified SRT textbooks and orthopedic training guidelines. Preceptors delivered two centralized lectures per month covering basic orthopedic theory and diagnostic and therapeutic specifications for common diseases, primarily in a didactic, one-way format. Practical teaching was fully integrated into daily clinical rotations. Trainees shadowed preceptors to complete outpatient consultations, medical record documentation, and surgical observation, and were permitted to perform only a small number of simple procedures under direct, full-time preceptor supervision. The entire preceptorship process was highly unstructured: preceptors typically provided ad hoc explanations based on cases encountered on the same day, and the content trainees learned was largely contingent on the clinical cases admitted that day, with no predefined systematic teaching curriculum. For assessment, this group underwent only one theoretical written test per month, one basic skills assessment per quarter, and one comprehensive mock examination prior to graduation. All assessment results were used solely as reference for trainees' graduation eligibility.

Following assessment, there was neither systematic feedback and problem analysis nor targeted tutoring and intensive training to address trainees' knowledge and skill deficiencies. No structured teaching feedback mechanism was established throughout the training period. Trainees could only ask questions opportunistically during clinical work when they had doubts, and mostly received immediate, informal oral responses. There was no proactive learning plan or continuous tracking preceptorship during the entire training cycle, resulting in an overall passive learning model characterized by "on-the-spot problem-solving as issues arise".

The optimized teaching group implemented a whole-process, closed-loop clinical teaching model aligned with the NMLE outline and final SRT assessment standards. Specific measures were as follows:

(1) Systematic theoretical instruction: Based on the examination outline and final assessment requirements, orthopedic content was divided into four modules: trauma, joint, spine, and hand and foot surgery. Core examination points for each module were curated by two senior preceptors. Two centralized module lectures were delivered per week, focusing on key and difficult concepts integrated with actual NMLE questions.

(2) Case-based learning (CBL): One CBL case discussion was held weekly. High-frequency NMLE cases (e.g., femoral neck fracture, lumbar disc herniation, complications after total knee arthroplasty) were selected. Trainees led case presentations and diagnostic and therapeutic reasoning analysis, while preceptors guided the development of management plans integrated with theoretical knowledge to strengthen theory-practice integration.

(3) Simulation skills training: Using orthopedic simulation operating tables, fracture models, and arthroscopic simulation equipment, one procedural skills training session was conducted weekly, covering fracture reduction, internal fixation placement, joint aspiration, and other core procedures. Preceptors standardized procedural techniques for each trainee, and a mock assessment was administered following training.

(4) Mentor accountability system and personalized learning: Each trainee was assigned two mentors with the rank of associate chief physician or higher. Personalized learning plans were developed based on trainees' baseline theoretical proficiency and learning capacity. One-on-one mentor consultations were conducted monthly to address learning difficulties in a targeted manner.

(5) Phased formative evaluation and feedback: Formative assessment was conducted monthly using the Objective Structured Clinical Examination (OSCE) format, covering four domains: theoretical Q&A, procedural skills, doctor-patient communication, and emergency management. Results were provided promptly after assessment, and targeted tutoring was delivered for identified weak areas. A mock NMLE was organized quarterly to analyze incorrect responses and identify weak knowledge modules. The core differences in weekly teaching arrangements between the two groups are shown in Supplementary Table 1.

Supplementary Table 1 Comparison of Weekly Teaching Arrangements between the Two Groups of SRT Trainees

Teaching module	Traditional teaching group	Optimized teaching group
Theoretical lecture	2 centralized lectures per month, mainly one-way teaching	2 modular lectures per week, explaining exam points combined with real licensing examination questions
Case discussion	No fixed arrangement, scattered explanation with clinical cases	1 CBL case discussion per week, trainees lead diagnosis and treatment analysis
Skills training	No fixed special training, only observation along with clinical operations	1 simulation skills training per week, preceptors standardize operation process for each trainee

Assessment and feedback	1 theoretical written test per month, no systematic feedback	1 OSCE formative assessment per month, immediate feedback after exam + targeted tutoring
Mentor communication	No fixed one-on-one communication, random Q&A	1 one-on-one communication with dual mentors per month, customized personalized learning plan

2.3. Assessment Indicators

The primary outcome indicator of this study was the first-attempt pass rate of the National Medical Licensing Examination, calculated as:

$$\text{First-attempt pass rate} = \frac{\text{Number of trainees who passed on the first attempt}}{\text{Total number of trainees in the group}} \times 100$$

The passing standard implemented the unified passing line set by the National Medical Examination Center.

Secondary outcome indicators were the final SRT assessment scores, including theoretical assessment score (full score 100 points, 60 points for passing), skill assessment score (full score 100 points, 60 points for passing), and total score (total score = theoretical score $\times$ 40% + skill score $\times$ 60%). All assessments implemented the unified standards for the final SRT assessment issued by the National Health Commission.

Process evaluation indicators included two aspects:

1 Teaching satisfaction: Evaluated using a well-validated SRT teaching satisfaction questionnaire (Cronbach's $\alpha=0.85$), covering 4 dimensions: teaching content, teaching methods, preceptors, and teaching arrangement. A 5-point Likert scale was adopted. The satisfaction rate was calculated as:

$$\text{Satisfaction rate} = \frac{\text{Number of "very satisfied" + "satisfied" responses}}{\text{Total number of trainees in the group}} \times 100$$

2 Clinical competence evaluation: Referencing the core competency framework for Chinese residents, an orthopedic SRT clinical competence evaluation scale (Cronbach's $\alpha=0.87$) was developed across four dimensions: theoretical knowledge, procedural skills, clinical reasoning, and doctor-patient communication, with a maximum score of 100 points. Evaluations were completed by three senior orthopedic chief physicians who were not involved in the teaching of this study. All evaluators underwent unified training and scale scoring calibration prior to assessment. A blinded design was employed during the evaluation process: evaluators were unaware of trainees' group assignment and scored independently based solely on trainees' on-site procedural performance and Q&A responses. The mean score of the three evaluators was used as the final score to control for observer bias. Grading criteria were: excellent (85–100 points), good (70–84 points), qualified (60–69 points), and unqualified (<60 points). The excellent rate was calculated as:

$$\text{Excellent rate} = \frac{\text{Number of excellent trainees}}{\text{Total number of trainees in the group}} \times 100$$

2.4. Data Collection and Statistical Methods

(1) Data collection: This study was a retrospective historical cohort study, and all data were retrospectively extracted from existing records. Specifically, baseline admission data, National Medical Licensing Examination scores, and final SRT assessment scores of the two groups were retrieved from the Education Department and the Orthopedic SRT Base Office (all were historical data formed during and after trainees' training). The teaching satisfaction questionnaires and clinical competence evaluation scales, which were uniformly distributed and collected on site by the SRT teaching secretary at the end of training, were traced. Questionnaires and scales with complete data were included in the analysis. All extracted data were independently entered and cross-checked by 2 researchers, and included in statistical analysis after verification, to ensure the authenticity, accuracy and completeness of data.

(2) Statistical analysis: Data analysis was performed using SPSS 26.0 statistical software (IBM Corp., Armonk, NY, USA). Continuous variables that followed a normal distribution as assessed by the Shapiro-Wilk test were expressed as mean $\pm$ standard deviation ($\bar{x} \pm s$) and compared between groups using the independent samples t-test.

Categorical variables were expressed as frequencies and percentages (%) and compared between groups using the chi-square test. Univariate chi-square analysis was conducted to identify factors associated with NMLE pass rates. Variables with a P-value < 0.1 in univariate analysis were entered into a multivariate unconditional logistic regression model. The regression analysis used "NMLE failure" as the positive outcome (coded as 1) to examine the effect of each factor on the risk of examination failure; an odds ratio (OR) < 1 indicated that the factor was a protective factor associated with reduced risk of examination failure. The significance level was set at $\alpha = 0.05$, and a two-tailed P-value < 0.05 was considered statistically significant.

3. Results

3.1. Comparison of First-Attempt Pass Rates of the National Medical Licensing Examination

Among the 45 trainees in the optimized teaching group, 40 passed the NMLE on the first attempt, yielding a first-attempt pass rate of 88.89%. Among the 42 trainees in the traditional teaching group, 29 passed the examination on the first attempt, corresponding to a first-attempt pass rate of 69.05%. The first-attempt pass rate was significantly higher in the optimized teaching group than in the traditional teaching group, with a statistically significant between-group difference ($\chi^2=5.21$, $P=0.022$) (Table 2).

Table 2. Comparison of First-Attempt Pass Rates of the National Medical Licensing Examination between the Two Groups

Group	Total number (n)	Number of passes (n)	Pass rate (%)	χ^2 value	P value
Traditional teaching group	42	29	69.05	5.21	0.022
Optimized teaching group	45	40	88.89	-	-

3.2. Comparison of Final SRT Assessment Scores

The theoretical score, skill score and total score of the final SRT assessment in the optimized teaching group were significantly higher than those in the traditional teaching group, with statistically significant differences (all $P<0.001$) (Table 3).

Table 3. Comparison of Final SRT Assessment Scores between the Two Groups (Points, $\bar{x} \pm s$)

Group	Total number (n)	Theoretical score	Skill score	Total score
Traditional teaching group	42	73.48 $\pm$ 7.12	76.25 $\pm$ 6.89	74.86 $\pm$ 6.53
Optimized teaching group	45	82.35 $\pm$ 6.21	85.62 $\pm$ 5.38	83.98 $\pm$ 5.76
t value	-	6.20	7.10	6.91
P value	-	<0.001	<0.001	<0.001

3.3. Comparison of Teaching Satisfaction and Clinical Competence Evaluation

The teaching satisfaction rate in the optimized teaching group was 95.56%, which was significantly higher than 78.57% in the traditional teaching group, with a statistically significant difference ($\chi^2=7.23$, $P=0.007$). The excellent rate of clinical competence evaluation in the optimized teaching group was 82.22%, which was significantly higher than 52.38% in the traditional teaching group, with a statistically significant difference ($\chi^2=9.45$, $P=0.002$) (Table 4).

Table 4. Comparison of Teaching Satisfaction and Clinical Competence Evaluation between the Two Groups (N, %)

Indicator	Traditional teaching group (n=42)	Optimized teaching group (n=45)	χ^2 value	P value
Teaching satisfaction	33 (78.57)	43 (95.56)	7.23	0.007
Excellent rate of clinical competence evaluation	22 (52.38)	37 (82.22)	9.45	0.002

Qualitative feedback from trainees in the optimized group was thematically analyzed, and representative comments included: "The weekly modular lectures clearly highlight examination points, making them far more efficient than self-study from textbooks"; "The simulation skills class allows repeated practice, so I feel much more confident when performing actual procedures"; "Mentors tailor learning plans to my specific weak areas, which is more effective than the indiscriminate practice I did before"; "The monthly OSCE assessment helps me identify procedures I haven't mastered early on, so I don't have to cram to fix gaps right before graduation".

3.4. Analysis of Influencing Factors for the Pass Rate

3.4.1. Univariate Analysis

Taking whether passing the National Medical Licensing Examination as the dependent variable, univariate chi-square analysis was performed on trainees' gender, age, education level, baseline score at admission, teaching mode, teaching satisfaction, and clinical competence grade. The results showed that baseline score at admission (≥ 70 points), optimized teaching mode, high teaching satisfaction, and excellent clinical competence were related factors affecting the pass rate ($P < 0.1$) (Table 5).

Table 5. Univariate Analysis of Influencing Factors for the Examination Pass Rate

Factor	Category	Total number (n)	Number of passes (n)	Pass rate (%)	χ^2 value	P value
Gender	Male	63	50	79.37	0.41	0.522
	Female	24	19	79.17	-	-
Age	<25 years	58	47	81.03	0.32	0.571
	≥ 25 years	29	22	75.86	-	-
Education level	Bachelor's	59	46	77.97	0.16	0.689
	Master's and above	28	23	82.14	-	-
Baseline score at admission	<70 points	41	28	68.29	5.12	0.024
	≥ 70 points	46	41	89.13	-	-
Teaching mode	Traditional teaching	42	29	69.05	5.87	0.015
	Optimized teaching	45	40	88.89	-	-
Teaching satisfaction	Dissatisfied/Average	11	5	45.45	7.86	0.005
	Satisfied/Very satisfied	76	64	84.21	-	-
Clinical competence evaluation	Non-excellent	48	33	68.75	8.94	0.003
	Excellent	39	36	92.31	-	-

3.4.2. Multivariate Logistic Regression Analysis

Taking failure in the National Medical Licensing Examination as the dependent variable, variables with $P < 0.1$ in univariate analysis (baseline score at admission, teaching mode, teaching satisfaction, clinical competence grade) were included as independent variables in the multivariate unconditional logistic regression model. The variable

assignment is shown in Table 6. The results demonstrated that the optimized teaching model, baseline score ≥ 70 points, and excellent clinical competence were independent protective factors against examination failure, each significantly associated with reduced risk of NMLE failure (all $P < 0.05$) (Table 7). The OR corresponding to the optimized teaching model was 0.21, indicating that after adjusting for other confounding factors, the risk of NMLE failure in the optimized teaching group was only 21% of that in the traditional teaching group.

Table 6. Variable Assignment Table

Variable	Assignment description
Dependent variable	Failure in the National Medical Licensing Examination: No (pass) = 0, Yes (fail) = 1
Baseline score at admission	< 70 points = 0, ≥ 70 points = 1
Teaching mode	Traditional teaching = 0, Optimized teaching = 1
Teaching satisfaction	Dissatisfied/Average = 0, Satisfied/Very satisfied = 1
Clinical competence evaluation	Non-excellent = 0, Excellent = 1

Table 7. Multivariate Logistic Regression Analysis of Influencing Factors for the Examination Pass Rate

Factor	β value	SE	Wald χ^2 value	P value	OR value	95% CI
Baseline score at admission (≥ 70 points)	1.42	0.59	5.79	0.016	4.14	1.30–13.17
Teaching mode (optimized teaching)	-1.56	0.57	7.45	0.006	0.21	0.07–0.64
Clinical competence evaluation (excellent)	1.38	0.61	5.13	0.024	3.97	1.20–13.13
Teaching satisfaction (satisfied/very satisfied)	1.05	0.65	2.61	0.106	2.86	0.80–10.20

4. Discussion

The findings of this study demonstrate that the optimized teaching model developed herein can effectively improve the performance of orthopedic SRT trainees in the two key examinations. Its value extends beyond the observed improvements in pass rates and scores; more importantly, it establishes a highly feasible practical pathway for implementing systematic, refined clinical teaching of orthopedics in municipal tertiary Class-A hospitals.

4.1. How Does Examination-Oriented Teaching Reconstruction Improve Pass Rates and Scores?

The first-attempt pass rate of the National Medical Licensing Examination in the optimized teaching group reached 88.89%, which was significantly higher than 69.05% in the control group. Further multivariate analysis also showed that the teaching mode itself was an independent protective factor against examination failure (OR=0.21, 95% CI: 0.07–0.64, $P=0.006$). This finding is consistent with the recent trend in medical education that emphasizes goal orientation and examination-driven learning [9].

The effectiveness of this mode lies in its realization of personalized management of the teaching process. First, teaching is modularly reconstructed strictly aligned with the examination outline, which solves the problem of disconnection between knowledge and assessment priorities in traditional teaching, and helps trainees build a systematic and targeted knowledge framework [10]. Second, frequent and regular CBL case discussions, based on constructivist learning theory, embed abstract theories into real clinical scenarios. They guide trainees to actively construct their diagnosis and treatment knowledge system instead of passively receiving information. This mode focuses on cultivating clinical decision-making ability, that is, the ability to quickly extract key information and form clear diagnosis and treatment logic when facing complex cases [11–13].

The whole-process formative evaluation (such as monthly OSCE), based on Bloom's mastery learning theory, constructs a dynamic learning monitoring network. It helps teachers timely detect trainees' weak links and carry out targeted interventions, avoiding the accumulation of knowledge and skill gaps before the examination [14]. Meanwhile, this simulated practice can also relieve trainees' examination anxiety and improve the stability of on-site performance. Finally, the dual-mentor system and personalized learning pathway constitute one of the core components underlying the model's effectiveness. Grounded in adult learning theory and self-determination theory, the efficacy of the mentor accountability system stems from three interrelated mechanisms: First, targeted feedback. Through continuous monitoring of trainees' learning progress, mentors can accurately identify knowledge and skill deficiencies and provide tailored improvement recommendations, substantially enhancing learning efficiency compared with traditional, generalized preceptorship. Second, autonomy support. Mentors engage trainees in the development of personalized learning plans, respect their individual learning pace and career development aspirations, fulfill the autonomy needs of adult learners, and foster intrinsic learning motivation. Third, emotional and psychological support. During SRT, trainees face the dual pressures of clinical responsibilities and examination preparation. The structured mentor consultation mechanism provides trainees with a reliable emotional outlet and professional confidence support, alleviates examination anxiety, and improves the stability of on-site performance. This arrangement provides robust support for trainees with weaker foundational knowledge and contributes to improved overall pass rates [15,16].

The impact of the systematic teaching mode is comprehensive and in-depth. In the final SRT assessment, the theoretical, skill and total scores of the optimized group were significantly better than those of the traditional group (all $P < 0.001$). This indicates that what we have improved is not only examination-taking ability, but also more fundamental clinical post competence. The advantage in theoretical scores directly stems from the highly consistent design of teaching content and final assessment standards. The breakthrough in skill scores is attributed to the standardized closed-loop of simulation training. This closed-loop can carry out targeted correction of operation details, and high-fidelity scenario simulation places trainees under the on-site pressure of rapid decision-making and teamwork, reversing the ability defect caused by the previous "emphasis on observation, neglect on practice" [17-18]. In addition, CBL case discussions and phased OSCE assessments comprehensively improve trainees' history taking, doctor-patient communication and clinical thinking levels through continuous practice. The superposition of these abilities is finally reflected in the overall improvement of total scores, which reflects the solid progress of trainees' overall clinical literacy.

4.2. Replication of the Mode: Advantages and Foundations for Implementation in Municipal Tertiary Hospitals

The optimized teaching mode has clear promotion value in orthopedics of municipal tertiary Class-A hospitals, and its feasibility is rooted in several practical foundations. As regional medical centers, municipal tertiary Class-A hospitals have abundant disease types and surgical volumes, which provide sufficient case sources for CBL teaching and clinical preceptorship [19-20], which is the premise for carrying out systematic teaching. Such hospitals usually have a stable team of senior physicians, who are fully capable of undertaking modular teaching and case preceptorship through targeted teaching method training. With reasonable teaching performance incentives, the enthusiasm of preceptors can be effectively stimulated [21-23]. Therefore, establishing a systematic teacher training and incentive system and incorporating teaching performance into evaluation is a key part of future promotion. The core of this mode is not to introduce expensive new equipment, but to standardize, systematize restructure and optimize the existing preceptorship process, which can be operated mainly relying on the existing clinical skills center and teaching management structure of the hospital. Meanwhile, the mode takes improving the pass rates of the "two key examinations" as the direct outcome, which is

consistent with the core indicators of tertiary hospital performance appraisal and SRT base evaluation [24], making it easy to obtain the understanding and support of hospital management, thus providing institutional guarantee for the implementation and sustainable operation of the mode.

4.3. From Teaching to Education: Building a Sustainable Management System

The deep value of the optimized mode lies in promoting a shift of educational focus from final evaluation to whole-process cultivation. Its core is to extend the focus from final assessment results to the whole process of physician growth. To realize this transformation, the key is to build a closely linked dynamic management system around all teaching links, making it an ecosystem that can be continuously optimized and self-regulated [25]. In departments with detailed subspecialty divisions such as orthopedics, this system requires specific and clear implementation paths. A practical starting point is to conduct reverse design based on the dual orientation of national core competency requirements for residents and final assessment standards, and develop standardized teaching plans covering all subspecialties and throughout the training process, to ensure that daily teaching is aligned with the final competency assessment goals from the very beginning.

A more breakthrough improvement is that this study transforms formative evaluation from a flexible reference indicator into a rigid teaching management tool. Specifically, process assessments such as monthly case evaluation and quarterly multi-station skill mock examinations need to be fixed with clear systems, and truly influence the adjustment of learning paths. These continuously generated assessment data can be used as archival records to draw a "learning trajectory map" of each trainee's ability fluctuation and shortcomings, and dynamically and timely adjust their personal learning plans accordingly. For example, when the system indicates that a trainee continues to perform poorly in the "spinal cord injury assessment" module, an early warning should be activated, and special tutoring resources should be coordinated for targeted intervention, so as to achieve precise support [26].

This refined process management inevitably requires a strong support system. The ideal configuration is to arrange dual mentors for each trainee: one focuses on long-term career development and clinical thinking cultivation, and the other goes deep into daily training, responsible for direct guidance and feedback of skill operations. At admission, a preliminary "competency profile" is formed through multi-dimensional baseline assessment, and a personalized learning plan is formulated accordingly. The plan is not static, but dynamically adjusted under the continuous tracking of mentors. For trainees with lagging learning progress, assistance should be launched quickly, with the core goal of diagnosing the root cause of learning difficulties and providing solutions rather than imposing penalties [27].

The fundamental purpose of these efforts to strengthen process management is not control, but guidance and empowerment. We hope that through this whole-process and supportive guidance, trainees will be promoted to complete the role transformation from passive recipients of knowledge to active constructors of their own knowledge system. Ultimately, we hope to cultivate not only medical students who can pass examinations, but also competent clinicians with continuous evolution ability who can cope with various clinical challenges in the future. This is essentially promoting a cultural shaping from "examination-oriented" to "application-focused", and even internalizing into a lifelong learning habit.

This study has several limitations that should be acknowledged. First, as a single-center historical cohort study with a limited sample size, the generalizability of our findings should be interpreted with caution. We adopted a historical control design and verified that the NMLE outline and difficulty remained generally stable between 2017 and 2022, with no major revisions. However, during the COVID-19 pandemic from 2020 to 2022, fluctuations in hospital clinical work rhythms, trainee rotation arrangements, and workload may have introduced confounding effects on our results. Future studies should

employ prospective designs to better control for confounding factors, and multi-center, large-sample prospective studies are needed to further validate the generalizability of our findings. Second, our study focused on short-term outcomes during SRT (scores of the "two key examinations" and teaching satisfaction). Long-term follow-up data are still lacking regarding whether this teaching model has a sustained positive impact on trainees' long-term clinical competence and career development. Additionally, we did not conduct subgroup analyses of teaching effects across different orthopedic subspecialties such as trauma, spine, and joint surgery; future work could develop more subspecialty-specific teaching protocols. There remains substantial scope for further optimization of the teaching model. For example, deeper integration of digital tools (e.g., online micro-lectures, virtual simulation surgery, intelligent question banks) could overcome temporal and spatial constraints and enhance teaching efficiency and trainee learning autonomy. How to quantitatively assess and improve trainees' implicit competencies such as clinical reasoning and humanistic care is also an important direction for future educational research.

References

1. National Health Commission of the People's Republic of China, Ministry of Education, and National Administration of Traditional Chinese Medicine, "Administrative Measures for Standardized Residency Training (Trial)," [Z], No. 49 [2014] of National Health and Family Planning Commission, 2014.
2. H. C. Ruan, L. J. Fan, X. B. Geng et al., "Analysis and reflection on the evaluation results of standardized residency training bases," *Chinese Journal of Medical Education*, vol. 41, no. 6, pp. 563–566, 2021.
3. General Office of the State Council, "Guiding Opinions of the General Office of the State Council on Accelerating the Innovation and Development of Medical Education," [EB/OL], No. 34 [2020] of the General Office of the State Council, Sep. 17, 2020. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/2020-09/23/content_5546373.htm?ivksa=1023197a. [Accessed: Mar. 29, 2026].
4. Orthopedics Branch of Chinese Medical Association, *Guidelines for Standardized Residency Training of Orthopedics (2020 Edition)*. Beijing: People's Medical Publishing House, 2020.
5. J. Zou and S. Liu, "Thoughts on standardized residency training in key orthopedic disciplines," *Education Progress*, vol. 12, no. 3, pp. 776–780, 2022.
6. S. C. Gao, A. L. Liang, W. Luo et al., "Thoughts on improving the quality of orthopedic rotation teaching in standardized residency training," *Chinese Journal of Medical Education Exploration*, vol. 15, no. 3, pp. 310–313, 2016.
7. L. Y. Yang, G. Ba, L. Q. Yang et al., "Problems and thoughts in standardized residency training of orthopedics," *China Continuing Medical Education*, vol. 9, no. 3, pp. 1–2, 2017.
8. L. Q. Shen, Y. H. Li, and J. Y. Chen, "Correlation between the results of the licensed physician qualification examination and resident graduation assessment," *Medical Education Management*, vol. 11, no. 5, pp. 614–619, 2025.
9. Y. Xu, H. W. Zhu, J. Chen et al., "Study on the effect and optimization of standardized residency training implementation from the perspective of satisfaction survey and Gibbs reflection cycle mode," *Chinese General Practice*, vol. 23, no. 5, pp. 867–870, 2025.
10. Talent Exchange Service Center of National Health Commission, "Outline of Theoretical Examination for Final Assessment of Standardized Residency Training (Trial)," [EB/OL], Dec. 30, 2024. [Online]. Available: <https://www.21wecan.com/wsrcw/c100224/202412/1002228.shtml>. [Accessed: Mar. 29, 2026].
11. W. Wei, C. C. Zhao, Y. Jiao et al., "Application of CBL-based tumor MDT outpatient teaching in standardized residency training of medical oncology," *Journal of Medical Science Yanbian University*, vol. 48, no. 11, pp. 128–132, 2025.
12. N. N. Kang, Y. Zheng, and R. Q. Zhang, "Preliminary exploration of CBL teaching method applied in standardized residency training of thoracic surgery," *Journal of Jiujiang University (Natural Science Edition)*, vol. 40, no. 3, pp. 113–117, 2025.
13. "Exploration on the application of centralized CBL mode in the first year of standardized general practitioner residency training," *China Continuing Medical Education*, vol. 13, no. 4, pp. 14–17, 2021.
14. W. B. Hua, Y. X. Du, F. W. Deng et al., "Application of competency-oriented Mini-CEX and DOPS in standardized orthopedic residency training," *Orthopaedics*, vol. 15, no. 3, pp. 253–257, 2024.
15. C. Ni, M. Li, B. Han et al., "Establishment and exploration of competency-oriented individualized residency training system," *Chinese Journal of Medical Education*, no. 5, pp. 771–774, 2015.
16. C. Y. Jiang, Y. L. Zhu, Y. L. Bai et al., "Construction and preliminary evaluation of clinical mentor system for standardized residency training of rehabilitation medicine," *Chinese Journal of Physical Medicine and Rehabilitation*, vol. 42, no. 5, pp. 460–463, 2020.
17. W. Wang, J. Z. Ding, and L. Xiao et al., "Application of scenario teaching mode combined with problem-based learning in orthopedic clinical teaching," *Beijing Medical Journal*, vol. 45, no. 2, pp. 181–184, 2023.

18. J. Yan, X. B. Wang, R. Hu et al., "Application of visualization technology in standardized residency training of anesthesiology," *Chinese Journal of Medical Education Exploration*, vol. 17, no. 8, pp. 839–843, 2018.
19. Z. Zhou, Y. Q. Ye, Z. Q. Chen et al., "Evaluation of teacher training effect in standardized residency training in a tertiary teaching hospital in Wuhan," *Medical Theory and Practice*, vol. 31, no. 10, pp. 1534–1536, 2018.
20. T. Li, "Evaluation on the implementation effect of standardized residency training: A case study of a tertiary hospital in Tangshan," M.S. thesis, North China University of Science and Technology, Tangshan, 2020.
21. J. Y. Li, H. B. Tao, S. Tao et al., "Discussion on the application of RBRVS in performance appraisal of public hospitals under cost control," *Chinese Hospital Management*, vol. 39, no. 10, pp. 46–48, 52, 2019.
22. Y. Z. Gao, R. Yang, and E. D. Huang, "Implementation path of public hospital performance management under the background of Healthy China — Based on comparison of public hospital performance management between Beijing and Shanghai," *Health Economics Research*, vol. 36, no. 7, pp. 17–20, 2019.
23. B. Wu, F. Chen, and P. L. Zhu et al., "Exploration and practice of performance appraisal reform in public tertiary hospitals," *Health Economics Research*, vol. 395, no. 3, pp. 68–70, 2020.
24. T. S. Xu, Y. Zheng, Q. F. Zhang et al., "Discussion on problems and countermeasures of teacher training for standardized residency training in Zhejiang Province based on training effectiveness — Taking the teacher training for standardized residency training held by Wenzhou Medical University as an example," *China Higher Medical Education*, no. 12, pp. 31–32, 2014.
25. J. Norcini, B. Anderson, V. Bollela et al., "Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference," *Medical Teacher*, vol. 33, no. 3, pp. 206–214, 2011.
26. W. C. McGaghie, S. B. Issenberg, J. H. Barsuk, and D. B. Wayne, "A critical review of simulation-based mastery learning with translational outcomes," *Medical Education*, vol. 48, no. 4, pp. 375–385, 2014.
27. T. Dornan, H. Boshuizen, N. King, and A. Scherpbier, "Experience-based learning: a model linking the processes and outcomes of medical students' workplace learning," *Medical Education*, vol. 41, no. 1, pp. 84–91, 2007.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.