

Pathways to improve the quality of clinical professional master's degree theses: a retrospective analysis of double-blind review results

Yijia Zhang

Chengdu Medical College

Linlin Wu

XXXXX

Zuobei Tang

XXXXX

Xinlu Li

XXXXX

Yanfeng Zhu

Chengdu Medical College

Fengmei Deng

Chengdu Medical College

Li Cheng

1000151@cmc.edu.cn

Chengdu Medical College

Research Article

Keywords: clinical professional master's degree, thesis quality, double-blind review, medical education, graduate education

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10999074/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Pathways to improve the quality of clinical professional master's degree theses: a retrospective analysis of double-blind review results

Yijia Zhang^{1†}, Linlin Wu^{2†}, Zuobei Tang², Xinlu Li³, Yanfeng Zhu¹, Fengmei Deng^{1*}, Li Cheng^{1*}

† These authors contributed equally to this work as the first authors.

* Corresponding author. E-mail: 1000151@cmc.edu.cn ; Tel: +86-028-62739177

Affiliations

1. School of Graduate Education, Chengdu Medical College, Chengdu 610000, P.R.China
2. School of Laboratory Medicine, Chengdu Medical College, Chengdu 610000, P.R.China
3. School of Basic Medical Sciences, Chengdu Medical College, Chengdu 610000, P.R.China

Abstract

Background: Thesis quality is a key indicator of graduate education quality. Double-blind review is an important quality-assurance procedure because it minimizes personal bias. This study analyzed double-blind review results for clinical professional master's degree theses at a provincial independent medical college from 2021 to 2023 to identify factors influencing thesis quality evaluation and to propose practical improvement pathways.

Methods: We retrospectively analyzed 1,156 double-blind review reports for 578 clinical professional master's degree theses submitted between 2021 and 2023. The sample included full-time and equivalency-based master's students; only initial expert review results were included. Descriptive statistics, chi-square tests, Kruskal–Wallis H tests, and text frequency analysis were used to examine differences by year, student type, supervisor academic rank, reviewer background, and clinical specialty. Review comments were analyzed qualitatively, and the top 20 problem-related keywords were compared across years.

Results: Median total scores were stable across years (2021: 80; 2022: 79; 2023: 80), but grade distributions differed significantly ($H = 6.880, p = 0.032$). The 2022 cohort had the lowest excellent rate (20.0%) and the highest failure rate (5.3%). Scores for theoretical and practical value, practical application, and thesis standardization declined significantly in 2023 relative to earlier years, whereas literature review and methodology and competence remained relatively stable. Student type was the strongest differentiator: equivalency-based students had a lower excellent rate (15.7% vs. 25.1%) and a higher failure rate (8.4% vs. 3.5%) than full-time students ($\chi^2 = 8.851, p = 0.031$). Supervisor academic rank, reviewer background, and clinical specialty were not significantly associated with review grades (all $p > 0.05$). High-frequency terms in review comments shifted from formal compliance toward scholarly contribution.

Conclusions: Thesis quality showed hidden differentiation and weakening of higher-order competencies despite stable median scores. Training institutions should (1) establish annual grade-distribution monitoring and early warning; (2) strengthen supervisor process management and guidance-time accountability; (3) improve double-blind review quality assurance; and (4) promote a shift from minimal compliance to substantive quality improvement.

Keywords: clinical professional master's degree; thesis quality; double-blind review; medical education; graduate education

Background:

The degree thesis serves as a comprehensive summary of the entire process of cultivating graduate students' overall competencies, and its quality is an important indicator and focal point for monitoring graduate education quality^[1–3]. In 2019, the Ministry of Education clearly stated in the *Notice on Further Standardizing and Strengthening Graduate Training Management* that all graduate training institutions must strictly emphasize thesis quality^[4]. Ensuring this quality has thus become a critical issue in the training of clinical professional master's degree students.

In recent years, a provincial independent medical college has adopted several measures to improve the quality of clinical professional master's degree theses. These measures include standardized format review, pre-defense, academic misconduct detection, double-blind review, and thesis defense. Among them, double-blind review is the one procedure in which the institution does not directly control the evaluation outcome. It is considered the most objective evaluation method because it minimizes bias from personal ties, and thus serves as a key step in ensuring thesis quality. Therefore, this study analyzed the double-blind review results of clinical professional master's degree theses at this institution from 2021 to 2023 to explore the content elements and influencing factors of thesis quality evaluation. Drawing on administrative experience, we also explored practical strategies for improving thesis quality. This study may provide a theoretical basis and practical reference for similar institutions.

Materials and Methods

Study Materials

In the double-blind review process, the names of graduate students and supervisors are concealed. Based on the research direction of each thesis, a third-party submission platform routes it to peer reviewers for evaluation. All reviewers hold associate senior or higher professional titles and are master's supervisors, and their personal information is likewise concealed. Each thesis is reviewed by two reviewers. Consequently none of the three parties—graduate student, supervisor, or peer reviewer—is aware of the others' identities during the entire review process, thereby ensuring objectivity throughout the process.

This study selected clinical professional master's degree theses submitted for double-blind review to a third-party platform at a provincial independent medical college between 2021 and 2023. The sample included theses from both full-time clinical professional master's students and candidates in the equivalent academic qualification pathway. A total of 1,156 review reports were included. Only the results from the initial expert review were included. The characteristics of the included theses are shown in Table 1.

Table 1: Statistics on Degree Dissertations (N=1,156)

Characteristic	Category	Number of review records (n)	Percentage (%)
A. Dissertation Source (Time & Program)			
Graduation Year	2021	268	23.2%
	2022	326	28.2%
	2023	562	48.6%
student's enrollment type	Full-time master's students	1073	92.8%
	Equivalency-based master's students	83	7.2%
B. Clinical & Academic Context			

Medical Specialty	Surgical Departments	275	23.7%
	Internal Medicine	232	20.2%
	Departments		
	Medical Imaging & Nuclear	132	11.5%
	Medicine		
	Anesthesiology	120	10.4%
	Obstetrics & Gynecology	77	6.7%
	Oncology	58	5.0%
	Pediatrics	52	4.5%
	Neurology	46	4.0%
	Dermatology &		
	Venereology	28	2.4%
	General Practice	26	2.2%
	Ophthalmology	24	2.1%
	Geriatrics	20	1.7%
	Rehabilitation Medicine &		
	Physical Therapy	20	1.6%
	Otorhinolaryngology	14	1.2%
	Laboratory Medicine	14	1.2%
	Emergency Medicine	10	0.9%
	Clinical Pathology	8	0.7%
Supervisor Academic Rank	Professor Physician (Senior)	615	53.2%
	Associate Professor/ (Sub-senior)	541	46.8%
C. Reviewer Characteristics			
Reviewer's Supervisor Status	Doctoral Supervisor	282	24.4%
	Non-Doctoral Supervisor	874	75.6%
Reviewer's Affiliation Tier	Project 985 university	164	14.2%
	Project 211 university	115	9.9%
	Other General Universities	877	75.9%

Review Criteria

The review criteria comprised five indicators: literature review (X1 = 10 points), theoretical and practical value (X2 = 15 points), methodology and competence (X3 = 30 points), practical application (X4 = 30 points), and thesis standardization (X5 = 15 points). The specific elements corresponding to each indicator are shown in Table 2. Reviewers scored each indicator based on the evaluation elements and summed the scores for a total . Based on the total score, theses were classified into four grades: excellent (≥ 85 points), good (75–84 points), pass (60–74 points), and fail (< 60 points). On the basis of the overall assessment, reviewers provided recommendations regarding whether a thesis could proceed to the defense stage: approve for direct defense (≥ 85 points), approve for defense after revision (60–84 points), or disapprove

for defense (<60 points).

Table 2: Evaluation Indicators and Evaluation Elements for Clinical Professional Master's Degree Theses

Evaluation Indicator	Evaluation Elements
Literature Review	Comprehensiveness and quality of summarizing domestic and international developments related to the selected topic.
Theoretical and Practical Value of the Topic	The cutting-edge nature and originality of the topic; relevance to real-world issues, ability to address practical problems, and reflection of the characteristics of the professional field.
Methodology and Competence	Appropriateness of research methods or design; scientific rigor in research procedures and processes; difficulty and workload of the research content; ability to comprehensively analyze, solve problems, and conduct investigations.
Practical Application	Practical value of the dissertation outcomes, ability to solve real-world problems, and the guiding or referential significance of the research findings.
Dissertation Standardization	Normative use of citations and terminological rigor; logical structure of the dissertation, accuracy and fluency of written expression.

Data Analysis Methods

Data were analyzed using SPSS 26.0 and NVivo 12. For continuous variables with a normal distribution, the mean $\pm$ standard deviation (SD) was used to describe central tendency when the data followed a normal distribution; otherwise, the median and interquartile range (IQR) were reported. The normality of the total thesis scores was assessed using the Kolmogorov–Smirnov test. Based on the distribution characteristics, intergroup comparisons were performed using independent samples t-tests (for normally distributed variables) or Kruskal–Wallis H tests (for non-normally distributed variables). Chi-square tests were applied to analyze the associations between categorical variables. All statistical tests were two-tailed, with a significance level set at $\alpha = 0.05$.

To supplement the quantitative findings, a qualitative content analysis was also conducted on the 1,156 double-blind review comments. The analysis followed four steps: (1) review comments were extracted to construct a raw corpus; (2) data preprocessing was performed, including removal of stop words, merging of synonyms, and standardization of professional terminology; (3) the top 20 high-frequency problem-related keywords were identified for each year from 2021 to 2023; and (4) cross-year comparisons were conducted to track changes in the reviewers' evaluative focus.

Results

Overall Scores of Clinical Professional Master's Degree Theses

As shown in Table 3, the median total scores of clinical professional master's degree theses from 2021 to 2023 were 80, 79, and 80, respectively, with similar interquartile ranges, indicating overall stability in score levels, despite some individual fluctuations..

Table 3. Median and Distribution of Overall Thesis Scores by Cohort (2021–2023)

Graduation Year	Number of Review Records (<i>n</i>)	Median Score	Interquartile Range (IQR)	Range (Min – Max)
2021	268	80	75.0 – 84.0	50.0 – 96.0
2022	326	79	74.0 – 83.0	37.0 – 94.0
2023	562	80	75.0 – 85.0	54.0 – 96.0
Total (2021–2023)	1,156	80	75.0 – 84.0	37.0 – 96.0

To better capture the distributional differences, we converted the continuous scores into grade categories. As shown in Table 4, significant cross-year differences emerged: among the 2022 theses, the "excellent" rate was the lowest and the "fail" rate the highest across the three years, while the distributions in 2021 and 2023 were relatively similar. The Kruskal–Wallis H test confirmed that these differences were statistically significant ($H = 6.880$, $p = 0.032$).

Table 4. Distribution of Clinical Professional Master's Dissertation Evaluation Grades by Year (2021–2023) [*n*(%)]

Submission Year	Excellent (≥ 85 points)	Good (75–84 points)	Fair (60–74 points)	Poor (< 60 points)	<i>H</i>	<i>P</i>
2021 (<i>n</i> =268)	65 (24.3)	146 (54.5)	48 (17.9)	9 (3.3)	6.88	0.032
2022 (<i>n</i> =325)	65 (20.0)	174 (53.5)	69 (21.2)	17 (5.3)		
2023 (<i>n</i> =563)	152 (27.0)	291 (51.7)	101 (17.9)	19 (3.4)		

Differences in the Distribution of Individual Thesis Evaluation Indicators

To further explore annual differences across specific evaluation indicators, we conducted a categorical statistical analysis of the scores for each indicator (Table 5). The score for “theoretical and practical value” was significantly higher in 2022 than in 2023 ($p < 0.05$). By contrast, the scores of “practical application” and “thesis standardization” declined markedly in 2023, significantly lower than those in 2021 and 2022 ($p < 0.05$). No significant annual changes were observed for the scores of “literature review” and “methodology and competence” ($p > 0.05$). These results suggest that the quality of theses at this institution exhibited a pattern of “stable foundational competencies but weakening higher-order competencies.” Graduate students performed relatively solidly in foundational abilities such as literature synthesis and the application of research methods, but showed notable deficiencies in higher-order abilities including the distillation of theoretical value, orientation toward clinical application, and adherence to academic writing standards.

Table 5. Differences in the distribution of various indicators in the thesis

Evaluation Dimension	Significant Pairwise Comparison (Year A > Year B)	Adjusted <i>p</i> -value
Theoretical and Practical Value of the Topic	2022 > 2023	$p < 0.05$
	2022 > 2023	$p < 0.05$
Practical Application	2021 > 2023	$p < 0.05$

Dissertation Standardization	2022 > 2023	$p < 0.05$
	2021 > 2023	$p < 0.05$
Literature Review	/	$p > 0.05$
Methodology and Competence	/	$p > 0.05$

Note: Post-hoc comparisons were performed using the Dunn-Bonferroni method following significant Kruskal-Wallis tests.

Effect of Student Type on Thesis Quality

To determine whether thesis quality varied by admission type, we compared the grade distributions of theses between full-time students and those in the equivalent academic qualification pathway. The results revealed a significant difference in grade distribution between the two groups ($\chi^2 = 8.851$, $df = 3$, $p = 0.031$), as shown in Table 6. Specifically, the proportion of theses rated as “excellent” was markedly higher among full-time students than among equivalent academic qualification students (25.1% vs. 15.7%), whereas the failure rate among equivalent academic qualification students was significantly higher than that among full-time students (8.4% vs. 3.5%). These findings indicate that admission type is a significant factor affecting the quality of clinical professional master's degree theses.

Table 6. Distribution of Dissertation Grades by Student Type [$n(\%)$]

Student Type	n	Excellent	Good	Fair	Poor
Full-time master's students	1073	269 (25.1)	568 (52.9)	198 (18.5)	38 (3.5)
Equivalency-based master's students	83	13 (15.7)	43 (51.8)	20 (24.1)	7 (8.4)
Total	1156	282 (24.4)	611 (52.9)	218 (18.9)	45 (3.8)

Effect of Supervisor's Professional Title and Peer Reviewer Background on Thesis Quality

To determine whether supervisors' professional titles and reviewers' backgrounds were associated with thesis evaluation outcomes, we conducted categorical statistical analyses of these variables. As shown in Table 7, none of the three variables—supervisors' professional titles, reviewers' institutional tier, or reviewers' doctoral supervisor status—had a significant association with review grades (all $p > 0.05$). Specifically, the proportion of theses rated “good or above” was 75.1% for those supervised by senior-title supervisors, compared with 79.7% for those supervised by associate senior-title supervisors, with no significant association ($\chi^2 = 3.383$, $p = 0.066$). For reviewers from “Double First-Class” universities, the proportion was 77.9%, versus 77.1% for reviewers from other universities, showing no statistically significant difference ($\chi^2 = 0.234$, $p = 0.89$). Similarly, the proportion rated “good or above” was 76.2% for reviewers with doctoral supervisor qualifications, compared with 77.6% for reviewers without such qualifications, again revealing no significant

difference ($\chi^2 = 0.216$, $p = 0.642$).

Table 7. Influence of Supervisor Academic Rank and Peer Reviewer Characteristics on Thesis Quality

Characteristic	Category	N (Reviews)	Good-and- above <i>n</i> (%)	Fair/Poor <i>n</i> (%)	χ^2 (<i>df</i>)	<i>p</i> -value
Supervisor Academic Rank	Senior (Professor/Chief Physician)	615	462 (75.1%)	153 (24.9%)	3.383	0.066
	Sub-senior (Associate Professor/Associate Chief Physician)	541	431 (79.7%)	110 (20.3%)		
	Double First-Class University (985/211)	279	217 (77.8%)	62 (22.2%)		
Reviewer Institutional Tier	Other General University	877	676 (77.1%)	201 (22.9%)	0.234	0.89
Reviewer Doctoral Supervisor Status	Doctoral Supervisor	282	215 (76.2%)	67 (23.8%)	0.216	0.642
	Non-Doctoral Supervisor	874	678 (77.6%)	196 (22.4%)		

Effect of Specialty on Thesis Quality

To determine the effect of different clinical specialties on thesis quality, we compared the review outcomes of theses across clinical specialties. To satisfy the assumptions of statistical analysis, specialties with fewer than 60 theses were combined into an “other” category; the grade distributions by specialty are shown in Table 8. The chi-square test revealed no significant difference in grade distribution among clinical specialties ($\chi^2 = 19.270$, $df = 15$, $p = 0.202$). Descriptively, the proportions of theses rated as “excellent” were relatively higher in internal medicine (28.5%) and the “other” specialty category (25.7%), whereas surgery had the lowest excellent rate (21.0%). Internal medicine and anesthesiology each had the highest proportion of “fail” grades (5.8% each), while the “other” category had the lowest failure rate (1.6%).

Table 8. Distribution of Dissertation Grades by Clinical Specialty [*n*(%)]

Specialty	<i>n</i>	Excellent	Good	Fair	Poor
Surgery	233	49 (21.0)	137 (58.8)	38 (16.3)	9 (3.9)
Internal Medicine	274	78 (28.5)	133 (48.5)	47 (17.2)	16 (5.8)
Obstetrics & Gynecology	77	14 (18.2)	46 (59.7)	15 (19.5)	2 (2.6)
Anesthesiology	120	29 (24.2)	58 (48.3)	26 (21.7)	7 (5.8)
Medical Imaging & Nuclear Medicine	133	30 (22.6)	70 (52.6)	27 (20.3)	6 (4.5)
Other Specialties	319	82 (25.7)	167 (52.4)	65 (20.4)	5 (1.6)
Total	1156	282 (24.4)	611 (52.9)	218 (18.9)	45 (3.8)

Note: The chi-square test showed no statistically significant difference in grade distribution among clinical specialties ($\chi^2 = 19.270$, $df = 15$, $p = 0.202$). To satisfy the assumptions of the statistical test, specialties with fewer than 60 cases (e.g., pediatrics, neurology) were merged into the “other” category.

Evolutionary Trends in High-Frequency Terms in Review Comments

To clarify the dynamic changes in reviewers’ evaluation concerns, this study conducted a high-frequency word analysis of 1,156 double-blind review comments from 2021 to 2023. The annual distribution trends clearly illustrate the evolution of reviewers’ evaluative focus, while also revealing recurring issues in theses across years, as shown in Figures 1a–1c.

In 2021, the high-frequency terms were concentrated in “logical coherence” (92 times), “references” (70 times), and “thesis writing” (68 times), which ranked the highest; “research findings” (66 times) and “domestic and international progress” (64 times) were also relatively frequent (Figure 1a). This distribution indicates that reviewers’ concerns in that year included both formal standards and substantive research content, though requirements for standardization still carried considerable weight.

In 2022, “logical coherence” (134 times) and “thesis writing” (120 times) remained the top two, with frequencies substantially higher than in 2021; “professional knowledge” (103 times) rose to third place, followed by “research findings” (95 times), “citation standards” (89 times), and “research methods” (83 times). “Innovativeness” ranked 15th (55 times), up from 18th (32 times) in 2021; “topic selection” rose in frequency (from 38 to 51 times), but its rank dropped from 14th to 17th, as other terms increased even more rapidly (Figure 1b). These rank changes suggest that reviewers’ focus was gradually shifting from formal compliance toward methodological rigor and professional competence—for example, “research methods” rose from outside the top 20 to 6th place, while “logical coherence” and “thesis writing” remained dominant but increased sharply in frequency.

In 2023, “logical coherence” (202 times) and “thesis writing” (191 times) remained the top two, while “innovativeness” (180 times) rose to third place, with “research findings” (177 times) and “domestic and international progress” (172 times) ranking fourth and fifth, respectively. Notably, the frequencies and ranks of “research methods” (147 times), “research limitations” (117 times), and “topic selection” (101 times) all increased significantly compared with the previous two years. Specifically, innovativeness rose from 18th place (32 times) in 2021 to 3rd place (180 times) in 2023. Meanwhile, topic selection rose from 17th in 2022 to 14th in 2023 (101 times), reflecting a marked increase in reviewers’ attention to the forward-looking nature and academic value of theses (Figure 1c). In addition, “research limitations” remained

around 11th–12th place, indicating a continuously strengthened demand for critical reflection in theses.

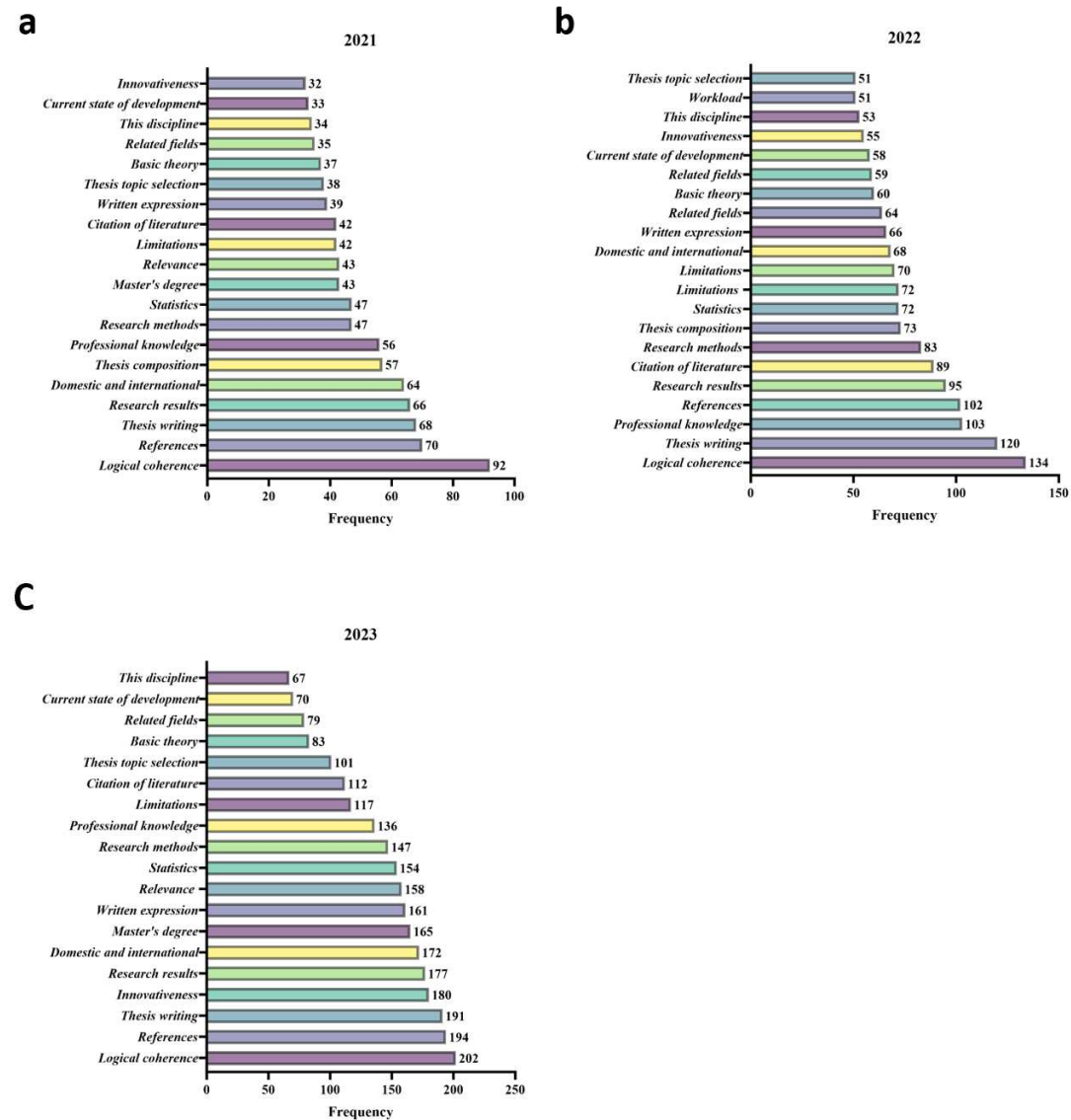


Figure 1. High-frequency problem-related keywords in double-blind review comments of clinical professional master's theses from 2021 to 2023. Bar charts show the top 20 keywords by frequency for (a) 2021, (b) 2022, and (c) 2023. The x-axis lists keywords, and the y-axis indicates frequency of occurrence. Keywords were extracted from 1,156 review comments after removing stop words, merging synonyms, and standardizing terminology. The evolution of keyword rankings reflects a shift in reviewers' focus from formal compliance toward scholarly contribution, with terms such as "innovativeness" and "topic selection" increasing in frequency over time.

Discussion

The Distribution of Thesis Quality Evaluation Exhibits a Skewing Phenomenon

This study examined the median total scores of clinical professional master's degree theses from 2021 to 2023. The medians were highly similar across years, and the interquartile ranges showed no notable differences, suggesting overall stability in thesis quality. However, converting continuous scores into grade categories revealed significant annual differences, indicating a dissociation between median-level stability and grade-level variability. A similar phenomenon was reported by Zhou et al. in

another local medical college^[5]. Central tendency indicators such as the median or mean are insufficient for capturing actual changes in thesis quality. When the excellent and failure rates fluctuate simultaneously, the stability of the middle segment may mask these fluctuations. Therefore, in monitoring graduate education quality, training institutions should complement routine average-score statistics with annual grade-distribution monitoring.

Thesis Quality Is Positively Correlated with Supervisor Guidance Time

Table 4 shows that the excellent rate in 2022 was the lowest (20.0%) and the failure rate the highest (5.2%) among the three years, with statistically significant differences compared with 2021 and 2023 ($p = 0.032$). Dimension-level analysis further revealed two distinct patterns: foundational competencies such as literature review and methodology and competence remained stable across years, whereas indicators heavily dependent on supervisor guidance and research time—including theoretical value, practical application, and thesis standardization—declined significantly in 2022. This pattern suggests that the observed quality decline in 2022 may be attributable, at least in part, to reduced supervisor guidance time and research constraints during that period.

The thesis research for the 2022 cohort was mainly conducted in 2021–2022, a period that coincided with a severe public health emergency. Clinical supervisors faced heavier clinical workloads and therefore had less time for guidance. Meanwhile, graduate students—especially those in the equivalent academic qualification pathway—had to reduce their research time due to front-line duties. This finding is consistent with Sahoo et al., who reported that 88% of medical postgraduates experienced thesis research disruption during the pandemic, with restricted supervisor guidance being the primary factor^[6]. The 2022 quality decline may be explained by the pandemic-induced “bidirectional time squeeze” on supervisors and students, suggesting a positive correlation between supervisor time and thesis quality—a pattern consistent with Zhang et al.^[7].

The Scholarly Contribution of Theses Remains Weak

Table 5 and Figures 1a–1c show that from 2021 to 2023, scores for “literature review” and “methodology and competence” remained relatively stable. In contrast, scores for dimensions of scholarly contribution—theoretical and practical value, practical application, and thesis standardization—declined significantly. This pattern was corroborated by the rising frequency of keywords related to scholarly contribution—such as “innovativeness,” “clinical value,” and “thesis standardization”—in review comments from 2021 to 2023, reflecting reviewers’ increasing expectations for substantive thesis quality. Overall, these results indicate that the scholarly contribution of clinical professional master's theses remains weak, a finding consistent with Ni et al.^[8].

Student Type Has a Significant Effect on Clinical Professional Master's Thesis Quality

Thesis quality varied by student admission type. Review results from this institution showed that equivalent academic qualification students produced theses of significantly lower quality than full-time students, mainly attributable to their in-service status, prominent work – study conflict, and weaker research foundations. Therefore, training institutions should adopt classified guidance strategies for different student groups. For this group, institutions should prioritize research methodology training, supervisor guidance, and thesis writing standards to narrow the quality gap with full-time students and promote more balanced thesis quality across student groups.

Supervisor's Professional Title and Reviewer Background Have No Effect on Thesis Quality

This study found no significant effect of supervisor's professional title or reviewer background on thesis quality, consistent with previous observations that supervisor rank does not necessarily reflect supervisory competence^[9]. Although supervisors with senior titles may have advantages in academic rank and research experience, their supervisory effectiveness is not necessarily superior to that of associate senior-title supervisors, possibly because clinical and administrative duties constrain their available time and energy. Therefore, improving thesis quality depends more on supervisors' substantive investment in guidance than on the advancement of their professional titles. Additionally, the finding that reviewer background did not influence evaluation outcomes further confirms that double-blind review can effectively minimize human bias, providing empirical support for the objectivity of the review system^[10].

Specialty Has No Effect on Clinical Professional Master's Thesis Quality

No statistically significant differences in thesis quality were found among clinical specialties, although descriptive data revealed notable disciplinary patterns. Internal medicine exhibited relatively high rates of both excellence and failure—a “dual-high” pattern suggesting considerable dispersion in training quality within this specialty, possibly related to large enrollment sizes and uneven supervisory loads. Liu et al. ^[11], in a multivariate regression of 1,275 clinical master's students, found that specialty had a limited influence on thesis quality. Anesthesiology had a failure rate comparable to internal medicine but a relatively low excellent rate, indicating that its thesis quality was generally concentrated in the middle-to-low range—a pattern consistent with Liang et al. ^[12], who also reported generally low quality in anesthesiology theses. The “other” category had the lowest failure rate; however, given its heterogeneous composition (formed by merging specialties with fewer than 60 cases), it cannot be used to infer the overall quality of those specialties. Therefore, in formulating quality assurance measures, institutions should focus on common factors—such as training process management, supervision practices, and research time protection—rather than on specialty as the primary distinguishing criterion.

Recommendations

Establish an Annual Quality Monitoring and Early Warning Mechanism Based on Grade Distribution

Current monitoring of thesis quality tends to rely on total scores or medians. These summary measures can reflect overall performance, but they are less useful for spotting changes at the two ends of the grade distribution, such as fewer excellent theses and more failures. When both occur, the median may remain unchanged, giving an impression of stability while quality differences are actually widening. Similar concerns have been raised in earlier work, which showed that total scores alone may not distinguish differences in overall quality among theses with similar scores ^[13].

One practical response is to monitor grade distribution each year, with particular attention to excellent and failure rates. A marked change in the shape of the distribution can act as an early signal, prompting a closer look at possible fluctuations in quality and providing evidence for adjusting training arrangements. Institutions may also consider a more integrated management process that covers proposal, mid-term review, pre-defense, double-blind review, defense, and post-defense review of low-scoring theses. This would allow problems to be identified earlier rather than only at the final pass/fail stage, and would support a shift from end-point inspection toward ongoing monitoring.

Establish a Supervisor Process Management Mechanism Based on Guidance Time

The current system for evaluating supervisors tends to emphasize measurable outputs, such as publications and grant funding, while paying less attention to how much time supervisors actually spend guiding their students. This study suggests that supervisory time is an important factor in thesis quality, and earlier research has shown that timely and focused guidance at key stages, such as topic selection, research design, and mid-term progress, helps prevent students from losing direction or falling behind schedule ^[7].

Drawing on our institution's experience, we propose several ways to strengthen process management for supervisors. One step would be to introduce formal supervisor management guidelines that set a minimum frequency for monthly meetings, provide a supervisor–student communication handbook, and require written records of guidance sessions. These records could be reviewed before graduation, adding a structural check on the input side. The handbook would also give students a clear basis for seeking guidance and encourage regular communication.

Another measure would be to set limits on the number of students a supervisor can take on, so that guidance quality is not diluted by excessive workload. For specific groups, such as clinical professional master's students in the equivalency pathway, guidance resources could be allocated more flexibly to balance supervisory loads. Finally, key training stages, including proposal, mid-term assessment, and pre-defense, could be organized centrally with all supervisors involved. This would help ensure that guidance remains continuous and effective. By tying supervisory responsibilities to required stages of training, supervision becomes a firm expectation rather than a voluntary extra, which would strengthen process-level review.

Establish a Process Assurance Mechanism for Double-Blind Review Quality

Thesis review is a critical link in quality assurance for graduate training. Multiple surveys have shown that the double-blind review system plays a positive role in curbing favoritism and ensuring academic fairness^[14]. However, its effectiveness depends on the standardization of the review process. Some universities have found in practice that relying solely on reviewers competence is insufficient to guarantee the quality of review comments, underscoring the need for a process assurance mechanism at the institutional level^[15].

Accordingly, training institutions can improve the process management of review quality in three aspects. First, establish a dynamic evaluation and renewal mechanism for reviewers: regularly assess their disciplinary match and the quality of their review comments, and replace those who provide perfunctory feedback, lack relevant expertise, or deviate from academic standards—thereby ensuring the academic competence of the reviewer pool. Second, clarify quality specifications for review comments: when sending theses for review, explicitly require that comments identify shortcomings and offer specific revision suggestions rather than superficial feedback^[14]; additionally, mandate detailed written comments or reasons for rejection on scoring forms, with electronic reports archived to ensure reviewers take each thesis seriously. Third, improve mechanisms for handling objections and conducting academic review within documents such as the implementation measures for thesis review. For example, this institution has established specific procedures for addressing objections to review comments in its master's thesis review measures, protecting the rights of degree applicants while also serving to inspect and supervise the review process.

Conclusions

This retrospective analysis of 1,156 double-blind review reports for 578 clinical professional master's degree theses from 2021 to 2023 showed that median total scores remained stable across years, but grade distributions differed significantly. The 2022 cohort had the lowest excellent rate and the highest failure rate, while dimensions reflecting higher-order competencies—theoretical and practical value, practical application, and thesis standardization—declined over time. Student type was the strongest factor associated with thesis quality: equivalency-based students had a lower excellent rate and a higher failure rate than full-time students. Supervisor academic rank, reviewer background, and clinical specialty were not significantly associated with review grades. Review comments increasingly emphasized scholarly contribution rather than formal compliance. These findings suggest that stable median scores may mask hidden differentiation in thesis quality and a weakening of higher-order competencies. To improve thesis quality, training institutions should establish annual grade-distribution monitoring and early warning, strengthen supervisor process management and guidance-time accountability, improve double-blind review quality assurance, and promote a shift from minimal compliance to substantive quality improvement. Future multi-institutional and longitudinal studies are needed to validate these pathways and evaluate their effectiveness.

List of abbreviation

ANOVA, Analysis of Variance

df, Degrees of Freedom

H, Kruskal–Wallis H statistic

IQR, Interquartile Range

NVivo, Qualitative Data Analysis Software

SD, Standard Deviation

SPSS, Statistical Package for the Social Sciences

χ^2 , Chi-square statistic

Statements and Declarations

Conflict of Interest Statement

No potential conflict of interest was reported by the authors.

Funding Statement

This work was supported by the research on factors influencing the quality of clinical professional master's degree theses and countermeasures based on blind review results (Grant No. YJG202310) from Chengdu Medical College, P.R. China; and the research on the quality management of graduate dissertations in the context of artificial intelligence (Grant No. 2024YB0405) from the Sichuan Provincial Society of Academic Degrees and Graduate Education, P.R. China.

Author Contribution Statement (CRediT)

1. Yijia Zhang: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft
2. Linlin Wu: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft
3. Zuobei Tang: Data Curation, Validation, Writing – Review & Editing
4. Xinlu Li: Investigation, Resources, Visualization
5. Yanfeng Zhu: Investigation, Resources, Visualization
6. Fengmei Deng: Supervision, Project Administration, Funding Acquisition, Writing – Review & Editing, Corresponding author
7. Li Cheng: Supervision, Project Administration, Funding Acquisition, Writing – Review & Editing, Corresponding author

Ethical Approval Statement

This study does not involve human participants or animal experiments, and no ethical

approval is required.

References

- [1] Li, M., & Chen, H. J. (2017). Typical Characteristics of Unqualified Academic Master's Degree Theses: Based on the Analysis of Expert Review Comments on Thesis Spot-Checking. *Academic Degrees & Graduate Education*, (6), 50-55.(in Chinese)
- [2] Jiang, G. Y. (2020). The empirical turn and realistic reflection of doctoral dissertation research in higher education: An analysis based on doctoral dissertations in higher education from 2010 to 2019. *China Higher Education Research*, (9), 66-72. (in Chinese)
- [3] Hao, M. J., Yang, Q., Zhang, C., et al. (2021). Review and prospect of academic degrees and graduate education in Chongqing over the past 40 years. *Academic Degrees & Graduate Education*, (9), 29-37. (in Chinese)
- [4] General Office of the Ministry of Education. (2019). Notice on Further Standardizing and Strengthening Graduate Training Management (No. Jiaoyuting [2019] 1). Retrieved from Ministry of Education website: http://www.moe.gov.cn/srcsite/A22/moe_826/201904/t20190412_377698.html(in Chinese)
- [5] Zhou, Y., Tan, J., & Huang, Y. D. (2024). Research on strategies for improving the quality of clinical medical master's professional degree theses in local medical colleges. *The Guide of Science & Education*, (23), 58-60. DOI:10.16400/j.cnki.kjdk.2024.23.019. (in Chinese)
- [6] Sahoo AK, Mishra N, Jain M, Sahoo M, Ghoshal P. Problems and solutions to conduct of thesis during the COVID-19 pandemic. *Korean Journal of Medical Education*. 2023;35(1):55–70.
- [7] Zhang, F., Hao, J. D., & Qu, G. H. (2025). A study on the impact of supervisor guidance frequency on the quality of postgraduate cultivation. *Journal of North University of China (Social Science Edition)*, 41(2), 10-18. (in Chinese)
- [8] Ni, H. N., Liu, J., & Shui, Y. P. (2024). Analysis of thesis quality among clinical medical master's degree professional postgraduates. *Health Vocational Education*, 42(7), 50-53. (in Chinese)
- [9] Cui, X. Y., Li, H., & Su, T. (2024). The influence of supervisor mentoring style on postgraduate innovation ability: A meta-analytic test. *Research on Graduate Education*, (5), 92-99. (in Chinese)
- [10] Pleskac TJ, Kyung EJ, Chapman GB, et al. Blinded versus unblinded review: A field study on the equity of peer-review processes[J]. *Management Science*. 2025.<https://doi.org/10.1287/mnsc.2022.01646>.
- [11] Liu Z H, Li L Y, Li H Y, et al.(2025) An analysis of factors influencing the dissertation quality of professional master's degree students in clinical medicine[J]. *Chinese Journal of Medical Education*, 45(11): 868-874. (in Chinese)
- [12] Liang, J. F., Hei, Z. Q., & Li, X. (2022). A bibliometric analysis of doctoral dissertations in anesthesiology in China. *Chinese Journal of Medical Education*, 42(10), 916-919. (in Chinese)
- [13] Peng, X. J., & Qu, J. S. (2024). A comprehensive quality grade identification method for doctoral dissertations and its validity verification. *Data Analysis and Knowledge Discovery*, (7), 164-175. (in Chinese)
- [14] Dong, Y. C., & Liu, J. (2025). Reflections on the Doctoral Dissertation Review Mechanism: Supplementing "Blind Review" with "Critical Examination." *Academic Degrees & Graduate Education*, (7), 78-86. (in Chinese)
- [15] Reziwanguli Abasi, Liu, D. Y., Wang, H. B., & Luo, L. (2026). Operational deviations and optimization pathways of the doctoral dissertation review system. *Chongqing Higher Education Research*, 14(3), 28-39. (in Chinese)