Journal of Clinical and Nursing Research, 2026, Volume 10, Issue 6
http://ojs.bbwpublisher.com/index.php/JCNR
Online ISSN: 2208-3693
Print ISSN: 2208-3685

Construction of an Evaluation Indicator System for Medical MOOCs Quality in China Using Delphi and Analytic Hierarchy Process

He Jiang[1], Yuan Li[1], Hui Chen[2], Chunyan He[2], Samdrup Phuntsok[2], Langtsu Tsering[2], Yinghui Liu[3]

[1]School of Nursing, Tianjin Medical University, Tianjin 300070, China
[2]Department of Medical and Nursing, Changdu Vocational and Technical School, Qamdo 854000, Xizang Autonomous Region, China
[3]Department of Nursing, Children's Hospital, Tianjin University /Tianjin Children's Hospital, Tianjin 300134, China

Abstract: *Objective:* This study addresses the extensive utilization of Massive Open Online Courses (MOOCs) in medical education and acknowledges the multitude of factors affecting their quality. Given the absence of a scientific and comprehensive assessment system, we aimed to formulate an evaluation indicator system for assessing the quality of medical MOOC construction in China and to determine the weight of each indicator. *Methods:* Using the Donabedian Structure-Process-Outcome model as theoretical guidance and based on relevant literature retrieval and analysis, we first constructed an evaluation indicator system for medical MOOC construction quality. Then, after two rounds of expert consultation, a final indicator content was built, and the weight of the indicators at each level was determined using the Analytic Hierarchy Process (AHP). *Results:* The effective recovery rate of the questionnaires in the two rounds of expert consultation was 68.75% (11/16) and 93.75% (15/16), respectively, and the expert authority coefficient (Cr) was 0.955 and 0.921, respectively. Kendall's W coefficients for indicators at the second and third levels were 0.174 and 0.238, respectively ($p < 0.01$), indicating substantial expert consensus, and the weighted values of the indicators were determined with a consistency ratio (C.R.) < 0.1. Consequently, the final evaluation index system comprises three first-level indicators, six secondary indicators, and 71 tertiary indicators. *Conclusion:* This study successfully established a comprehensive evaluation indicator system for assessing the quality of medical MOOCs, which is characterized by scientific rigor and reliability, and could serve as a valuable reference for evaluating the quality of Medical MOOCs in China.

Keywords: MOOC; Medical education; Indicator system; Delphi method; Analytic hierarchy process

Online publication: Jun 30, 2026

1. Introduction

Massive Open Online Courses (MOOCs) have emerged as a novel form of open online education, drawing increasing attention from governmental bodies, educational institutions, and educators since their inception. MOOCs possess distinct advantages when compared to traditional online education, including a wealth of high-quality educational resources, open accessibility, and ease of use [1,2]. Additionally, the COVID-19 pandemic outbreak has stimulated increasing interest and exploration into the possibilities and potential of online education [3]. Notably, in recent years, there has been a significant use of MOOCs. For example, China's MOOC initiative, initiated by universities in 2013, has achieved global leadership and ranks first in regard to scale and adoption over the past decade [4–6].

At present, universities and colleges worldwide are actively launching programs and courses that incorporate MOOCs, thereby indicating that the utilization of mobile and digital technology holds immense significance in medical education. MOOCs have found widespread application in healthcare disciplines, including clinical medicine education. Thus, MOOCs have emerged as a crucial online teaching method, facilitating the sharing of high-quality educational resources and promoting equitable education. However, the growing use of MOOCs has raised questions about their teaching quality. MOOCs have inherent limitations and controversies, including high dropout rates, low completion rates, a lack of standardized criteria for evaluating their quality, and challenges in ensuring effective learning outcomes and teaching quality [7,8]. Overcoming these obstacles is essential for MOOCs to become a dominant educational method in the future, with the quality of course construction playing a central role. As widely acknowledged, multiple factors influence the quality of medical MOOCs, encompassing elements such as the platform, course management, and interactions between teachers and students [9]. Consequently, educators are increasingly conducting research to comprehensively assess course quality and improve it. Indeed, course quality is the linchpin of all education, and any educational format must ensure its quality to achieve long-term viability. Therefore, it is imperative to address the challenge of scientifically and effectively evaluating the educational quality of MOOCs and developing a corresponding evaluation system. This issue holds significant value for all stakeholders, including teachers, students, and MOOC platforms. Medical MOOCs have their distinctive characteristics, requiring course planners and designers to align these curricula with the most current trends in medical developments and clinical environments [10]. Despite existing literature highlighting the pivotal role of quality in MOOCs, there has been relatively limited research on how to conduct a scientific and comprehensive assessment of the quality of medical MOOCs.

Present literature reveals that the evaluation of China's medical MOOCs predominantly follows a summative approach, often neglecting the importance of process evaluation [11]. Some studies highlight the necessity of strong instructional design as a fundamental quality requirement for ensuring the effectiveness of medical MOOCs, while others have focused on elements that affect the quality of medical MOOCs, such as platform design, instructors, and students [1,4,8,9,12,13].

However, it is important to note that there remains a lack of quality index systems or standardized evaluation tools for conducting comprehensive assessments of medical MOOC construction in China [11]. Therefore, it is important to establish a scientific and comprehensive quality indicator system for medical MOOCs, combining expert insights with rigorous statistical methods, as such a system could serve as valuable guidance for the advancement of medical MOOC teaching in China.

2. Methods

The development of the evaluation index system for assessing the quality of medical MOOC construction occurs through three key stages: first, the creation of the initial set of indicator items; second, the establishment and refinement of the indicator system through a Delphi consensus process; and third, the validation of reliability via weight analysis.

2.1. Design

The Delphi technique is a systematic communication approach involving multiple rounds of questionnaires with a panel of experts to achieve consensus [14]. In this research, the Delphi technique was used to analyze preliminary evidence-based indicators from literature through two rounds of expert consultation, such that the opinions of experts were consistent, and finally, the quality indicators were developed. The three-dimensional (Structure-Process-Outcome) quality structure model serves as the theoretical foundation for the development of quality evaluation indicators that explain how structure and process indicators influence outcomes [15]. Using the "three-dimensional quality structure" model as a guiding framework, a preliminary evaluation indicator system for assessing the quality of medical MOOC construction was established based on a relevant literature review and analysis. Then, the Analytic Hierarchy Process (AHP) was used to determine the weights assigned to indicators at all levels.

2.2. Establish a study group

The study group comprises nine members, including two individuals with senior professional titles, three with associate senior professional titles, two with intermediate professional titles, and two medical graduate students. Each member possessed substantial knowledge about the study subjects and was responsible for key tasks, such as literature retrieval, letter inquiry preparation, expert communication and consultation, data collection, and statistical analysis.

The primary roles of the study group included the discussion and assessment of evidence-based indicators, the creation of a preliminary draft of essential indicators, the design of an expert inquiry form, and the selection of members for the Delphi expert study group. The core mission involved assembling specialists with a comprehensive understanding of medical MOOC construction quality and eliciting their insights and opinions regarding the indicator system developed in this study.

2.3. Generation of the first draft

A systematic literature search was performed using CNKI, Wanfang, VIP, PubMed, and Medline databases from the inception of each database using the following main search terms: "MOOC construction quality", "indicator system", "Delphi method", and "Analytic Hierarchy Process (AHP)". Based on the Donabedian three-dimensional quality evaluation model and combined with the literature on the evaluation of MOOC teaching quality, Online teaching quality, and Medical curriculum quality, the research group developed a preliminary draft of the evaluation indicator system, which consisted of 3 first-level indicators, 6 second-level indicators, and 72 third-level indicators. The first draft was verified for readability and feasibility by two medical education experts.

2.4. Delphi process

2.4.1. Confirmation of the expert panel and inclusion criteria

This study selected 15 medical education specialists from universities in 11 cities, including Beijing, Shanghai, Tianjin, Xi'an, etc., utilizing the purposeful sampling method. Inclusion criteria comprised individuals holding an associate senior professional title, possessing a Master's degree or higher, having a minimum of 10 years of experience in medical teaching at a university, involvement in the development of at least one medical MOOC course, granting informed consent, demonstrating active engagement in this study, and a commitment to completing two rounds of questionnaires.

The panel of experts in this study, as summarized in **Table 1**, consisted of individuals aged 40 to 55, with extensive experience in medical teaching ranging from 12 to 39 years. Seven experts held senior professional titles, while five held associate senior professional titles. Among the panel, 12 had obtained Doctoral degrees, and three held Master's degrees.

Table 1. Experts' characteristics

Pseudonym	Gender	Age (years)	Years of teaching in medicine	Regions
E1	Man	41	12	Beijing
E2	Woman	46	25	Shanghai
E3	Man	51	17	Shanghai
E4	Man	40	17	Tianjin
E5	Man	44	18	Wuhan
E6	Man	55	21	Chengdu
E7	Woman	52	17	Nanjing
E8	Woman	46	14	Changsha
E9	Woman	54	39	Changsha
E10	Man	54	25	Changsha
E11	Woman	45	13	Xiamen
E12	Man	48	24	Hangzhou
E13	Woman	55	17	Taiyuan
E14	Woman	50	14	Xi'an
E15	Woman	48	25	Zunyi

The role of the expert panel in this study was to offer insights and recommendations concerning the content of the two consultation rounds and to evaluate the suitability of each aspect independently. The assessments provided by expert panel members were submitted anonymously, ensuring confidentiality. Their participation was entirely voluntary, and they provided informed consent without receiving any compensation. To finalize the evaluation system, the research group integrated the various categories and components based on the expert opinions. Consequently, the authority of the final evaluation system was determined based on the input, opinions, and recommendations of the expert panel.

2.4.2. Questionnaire development

The research questionnaire was structured into three sections: an introductory section, the core content of the questionnaire, and the experts' basic information. Firstly, the questionnaire's introduction provided context, elucidating the study's background, purpose, and significance. Secondly, the main body of the questionnaire presented each evaluation indicator along with a method for assessing the importance of each item. The experts were tasked with rating each item on a Likert scale ranging from 1 to 5, where 5 represented the highest significance. Additionally, they were encouraged to offer comments and propose additional items. Lastly, the experts' basic information encompassed details such as age, educational background, years of experience in medical education, professional position, professional title, familiarity with the indicators (Cs), and the rationale underpinning their judgments (Ca).

2.4.3. Expert consultation procedure

Two rounds of Delphi consultation were conducted via e-mail, and the experts were given a two-week window to respond after receiving the questionnaire. In the first round, the results were collected, numbered, sorted, and categorized for subsequent discussion. The questionnaire for the second round was designed through group discussions and revisions, taking into account relevant literature and standards. After two rounds of consultation, consensus was achieved among the experts, concluding the consultation process.

2.5. Research and ethics approval

Ethical approval for this study was granted by the Ethics Committee of Tianjin Medical University (Reference Number: TMuhMEC2023062). All expert participants were anonymized throughout the research process, and data were collected solely for scholarly purposes. Participants were formally invited via email, and explicit informed consent was obtained from each contributor, ensuring their voluntary participation and understanding of the study's objectives.

2.6. Data analysis

Data analysis was conducted using SPSS version 22.0, with the response rate of the questionnaire used as a metric to assess the level of expert engagement. The degree of authority among experts was determined using the authority coefficient (Cr), which takes into account the expert's educational background, the basis for their judgments (Ca), and their familiarity with the indicators (Cs). Cr was computed using the formula $Cr = (Ca + Cs) / 2$, and the results were considered acceptable when Cr exceeded 0.7, as previously described [16]. The degree of expert opinion concentration was evaluated using the mean and standard deviation of item importance scores. Consistency in expert opinions was assessed through the coefficient of variation (CV). A consensus was deemed to exist when the mean importance score reached 3.5 or higher, the coefficient of variance was less than 25%, and at least 80% of experts gave an item 5 points. Each indicator's weight was determined using AHP, and the geometric mean method was chosen to calculate the weight value. In regard to consistency testing, consistency index (C.I.) and average random consistency index (R.I.) were used, and the ratio of C.I. and R.I. was considered the consistency ratio (C.R.) [17]. A judgment matrix was considered to have a satisfactory weight distribution when the C.R. was less than 0.10, indicating satisfactory consistency.

3. Results

3.1. Enthusiasm and authority of experts

In this study, we conducted two rounds of expert consultation with 15 experts from 12 major cities, including Beijing, Shanghai, Tianjin, Wuhan, Chengdu, Nanjing, Changsha, Xiamen, Hangzhou, Taiyuan, and Xi'an. The first round of the Delphi process yielded 11 complete responses, resulting in a response rate of 73.33%, while the second round involved 15 expert panelists, achieving a recovery rate of 100.00%. The Ca values for the two rounds were 0.927 and 0.942, with Cs values of 0.982 and 0.900. The resulting Cr values were 0.955 and 0.921, indicating a high level of expert authority, in line with established criteria.

3.2. Degree of coordination and concentration

Kendall's W and the chi-square test were employed to assess the degree of coordination among expert opinions. Kendall's W yields values within the range of 0 to 1, with higher values indicating greater consensus and improved item coordination. Due to the limited number of first-level indicators (fewer than 7), only second and third-level indicators were analyzed. In the first round of expert consultation, Kendall's W values for the second and third levels were 0.143 and 0.240, respectively. In the second round, these values were 0.174 and 0.238. Additionally, the chi-square test indicated statistically significant differences ($p < 0.05$), indicating that the results were acceptable [16].

Next, the importance value and coefficient of variance were used to analyze the degree of concentration of expert opinions. The results showed that the degree of expert opinion concentration was better for higher importance assignments and smaller CV values. In this study, expert consultation resulted in mean importance scores ranging from 3.64 to 5.00 for each indicator and CV values ranging from 0 to 0.245, indicating a relatively high level of consensus and general agreement among the experts [16].

3.3. Results of expert consultation

In the first round, we observed a 90.3% expert agreement rate, with 6 experts making 20 recommendations, among which 7 were accepted (adoption rate, 35.0%). In the second round, the agreement rate increased to 94.4%, with 7 experts providing 35 recommendations (adoption rate, 34.3%). The final indicator system comprises 3 first-level indicators, 6 second-level indicators, and 71 third-level indicators. Additional details of the expert consultation can be found in **Table 2**.

Table 2. The results of the two rounds of expert consultation are summarized

	Structure		**Process**		**Outcome**		**All**	
	Round 1	**Round 2**	**Round 1**	**Round 2**	**Round 1**	**Round 2**	**Round 1**	**Round 2**
Second-level	2	2	3	3	1	1	6	6
Third-level	30	26	34	36	8	9	72	71
Expert opinions	10	13	9	18	1	4	20	35
Amendments	7	9	4	9	1	0	12	18
Additions	1	3	1	9	0	4	2	16
Deletions	2	1	4	0	0	0	6	1
Adopt opinions	6	5	1	6	0	1	7	12
Adoption rate (%)	60.0	38.5	11.1	33.3	0	25.0	35.0	34.3

3.4.Weight distribution of the quality evaluation system

According to the principles of AHP, the evaluation index system for medical MOOC construction quality was organized into 4 hierarchical levels. We constructed multiple judgment matrices to assess the consistency of each first-level indicator, and the findings revealed that all CI and CR values were less than 0.1. As shown in **Table 3**, this confirms the absence of logical errors in the determination of first-level indicator weights and highlights the consistency of judgments. Furthermore, we derived the weight coefficients for each indicator, which are presented in **Table 4**.

Table 3. The result of the consistency test of the second-level indicators

Indicators	CI	CR
Platform technique	0.030	0.020
Curriculum management	0.057	0.036
Team of teachers	0.073	0.047
Course content	0.045	0.029
Student participation	0.046	0.032
Student achievement	0.023	0.016

Table 4. Importance of assignment and weight of the medical MOOC construction quality evaluation system

Indicators content	Mean (SD)	CV	Weight	Combination weight
1 Structure			0.250	-
1.1 Platform technique	4.40 ± 0.63	0.144	0.250	0.063
1.1.1 The MOOC platform interface style is unified, concise, and elegant	4.13 ± 0.92	0.222	0.046	0.003
1.1.2 The MOOC platform interface has a navigation design that is straightforward to use	4.20 ± 0.68	0.161	0.061	0.004
1.1.3 The MOOC platform offers operating instructions and answers to common questions	4.00 ± 0.63	0.158	0.033	0.002
1.1.4 The video image is steady and clear, the composition is appropriate, and the sound quality is excellent	4.73 ± 0.46	0.097	0.208	0.013
1.1.5 The video has accurate subtitles	4.33 ± 0.62	0.142	0.084	0.005
1.1.6 Multiple play functions are available in the video, such as full-screen, double-speed, pause, and repeat	4.53 ± 0.83	0.184	0.145	0.009
1.1.7 The video has the function of keeping records of the keywords used, facilitating medical students to retrieve related information	4.13 ± 0.83	0.202	0.046	0.003
1.1.8 A search function is available on the course platform to assist medical students in finding learning resources	4.13 ± 0.99	0.240	0.046	0.003
1.1.9 The MOOC platform has a good operation response and can meet the needs of a large number of medical students to visit at the same time	4.73 ± 0.46	0.097	0.208	0.013
1.1.10 The MOOC platform has extensive administration capabilities and can monitor and analyze the learning process	4.50 ± 0.84	0.186	0.123	0.008
1.2 Curriculum management	4.73 ± 0.46	0.097	0.750	0.173
1.2.1 In curricular development, there is a complete testing procedure	4.40 ± 0.74	0.168	0.079	0.014
1.2.2 Invite medical students to participate in the MOOC's development and testing	4.13 ± 0.74	0.180	0.037	0.006
1.2.3 MOOC construction and administration are aided by adequate funding and personnel	4.87 ± 0.35	0.072	0.175	0.030
1.2.4 The course group holds regular teaching meetings	4.33 ± 0.62	0.142	0.061	0.011
1.2.5 The duration of the MOOC is reasonable, typically from 4 to 8 weeks	3.87 ± 0.92	0.236	0.020	0.004

1.2.6 Selection of those with high academic attainments and rich teaching experience to work as course lecturers	4.47 ± 0.64	0.143	0.093	0.016
1.2.7 There are periodic tests in the course	4.20 ± 0.94	0.224	0.048	0.008
1.2.8 There are teaching activities interspersed in the video, such as quizzes, surveys, reflection questions, and so on	3.64 ± 0.75	0.205	0.014	0.002
1.2.9 The course provides expanded learning resources, such as texts, books, links, etc.	3.93 ± 0.83	0.211	0.025	0.004
1.2.10 Course notifications are sent in various ways, including the platform, e-mail, and SMS	4.20 ± 0.78	0.185	0.048	0.008
1.2.11 Teachers and students have established interactive discussion forums, such as QQ, WeChat, etc.	3.73 ± 0.84	0.237	0.017	0.003
1.2.12 The course includes peer or self-evaluation as well as comprehensive grading standards	3.93 ± 0.96	0.245	0.025	0.004
1.2.13 The assessment of the course grade is reasonable and proportionate	4.13 ± 0.64	0.155	0.037	0.006
1.2.14 Certificates or degrees are provided according to the process evaluation and the final evaluation	4.33 ± 0.72	0.167	0.061	0.011
1.2.15 The course is periodically evaluated and revised to reflect the most recent developments in medical knowledge	4.60 ± 0.51	0.110	0.112	0.019
1.2.16 The feedback from students can be used to review or improve the course	4.73 ± 0.46	0.097	0.149	0.026
2 Process			0.644	-
2.1 Teaching staff	4.80 ± 0.41	0.086	0.297	0.219
2.1.1 The teacher faces the camera naturally	4.40 ± 0.74	0.168	0.037	0.008
2.1.2 The language used by the teacher is standardized and clear, with a suitable speaking tone and speed	4.47 ± 0.74	0.166	0.051	0.011
2.1.3 Teachers can effectively organize online discussions and interactions	4.47 ± 0.64	0.143	0.051	0.011
2.1.4 The teachers' teaching schedule is reasonable, and the teaching tasks can be completed on time	4.40 ± 0.51	0.115	0.038	0.008
2.1.5 The teaching method matches the content of MOOCs	4.80 ± 0.41	0.086	0.125	0.027
2.1.6 The organization of the MOOC content is logical, and the knowledge point arrangement is reasonable	4.93 ± 0.26	0.052	0.148	0.032
2.1.7 The medical knowledge taught by the teachers is up-to-date, cutting-edge, and practically applicable	4.47 ± 0.52	0.115	0.051	0.011
2.1.8 The medical knowledge taught by the teachers can meet students' learning needs	5.00 ± 0.00	0.000	0.184	0.040
2.1.9 Teachers focus on the intersection of knowledge in different subjects	4.60 ± 0.51	0.097	0.077	0.017
2.1.10 Teachers attach importance to imparting knowledge and educating people	4.73 ± 0.46	0.086	0.105	0.023
2.1.11 Teachers reply and organize valuable posts in a timely manner	4.27 ± 0.70	0.165	0.034	0.007
2.1.12 Teachers promptly correct students' papers, provide feedback, and answer inquiries	4.20 ± 0.78	0.185	0.024	0.005
2.1.13 Teachers organize MOOC-related offline activities	3.93 ± 0.96	0.245	0.016	0.004
2.1.14 The teacher's assessment of students is clear, fair, and effective	4.53 ± 0.74	0.164	0.058	0.013
2.2 Course content	4.93 ± 0.26	0.052	0.540	0.331

2.2.1 The course provides an introduction to the teachers or teaching team	4.60 ± 0.63	0.137	0.083	0.028
2.2.2 The course introduces the primary teaching content and objectives	4.67 ± 0.49	0.104	0.105	0.035
2.2.3 The assessment method and credit certification are clearly stated in the course	4.87 ± 0.35	0.072	0.165	0.054
2.2.4 The chapter's learning objectives are listed in the course and are measurable	4.53 ± 0.74	0.164	0.067	0.022
2.2.5 The video length of each knowledge point is appropriate, ranging from 5 to 10 minutes	4.40 ± 0.63	0.144	0.045	0.015
2.2.6 The course's extended learning materials are accessible and can meet students' learning needs	4.33 ± 0.62	0.142	0.033	0.011
2.2.7 Detailed rules are set to tutor student participation in course discussions.	4.47 ± 0.64	0.143	0.059	0.019
2.2.8 The course provides detailed assignment and examination requirements, such as starting and ending times	4.80 ± 0.41	0.086	0.128	0.042
2.2.9 The course can help students develop their cognitive abilities, experimental operation abilities, and clinical thinking abilities	4.47 ± 0.52	0.115	0.057	0.019
2.2.10 The course content can assist students in enhancing their self-learning ability	4.87 ± 0.35	0.072	0.160	0.054
2.2.11 The course content can assist students in enhancing their ability to cooperate	4.33 ± 0.52	0.119	0.033	0.011
2.2.12 The course content and delivery take into account the diversity of cultural values	4.40 ± 0.74	0.168	0.045	0.015
2.2.13 Regardless of cultural background, everyone can understand the examples in the course	4.07 ± 0.70	0.173	0.020	0.007
2.3 Student participation	4.60 ± 0.63	0.137	0.163	0.107
2.3.1 Students can watch MOOCs in full and on time	4.67 ± 0.62	0.132	0.201	0.022
2.3.2 Students watch the MOOCs with an earnest and attentive attitude	4.33 ± 0.82	0.188	0.061	0.007
2.3.3 Students can complete homework and unit tests on time	4.60 ± 0.63	0.137	0.141	0.015
2.3.4 Students completed homework and unit tests with good quality	4.60 ± 0.51	0.110	0.164	0.018
2.3.5 Students can carefully review the courseware or extended learning materials	4.47 ± 0.83	0.187	0.104	0.011
2.3.6 Students can evaluate each other based on the real situation of their classmates	4.40 ± 0.74	0.168	0.078	0.008
2.3.7 While viewing MOOCs, students can actively engage with teachers or classmates online	4.20 ± 0.86	0.205	0.048	0.005
2.3.8 Students can participate actively in the interaction of posts on the course platform	4.67 ± 0.49	0.104	0.172	0.018
2.3.9 Students can actively participate in off-platform discussion groups (such as QQ, WeChat, etc.)	4.07 ± 0.80	0.196	0.033	0.004
3 Outcome			0.107	-
3.1 Student achievement	4.60 ± 0.51	0.110	1.000	0.107
3.1.1 The final grade of the course (background data: the score of students)	4.53 ± 0.64	0.167	0.061	0.007
3.1.2 Course completion rate (the percentage of students who finished the MOOC of all the students who registered in it)	4.33 ± 0.72	0.086	0.041	0.004
3.1.3 Students master the knowledge of the course	4.80 ± 0.41	0.093	0.194	0.021

3.1.4 The students' clinical thinking ability has been improved	4.80 ± 0.45	0.097	0.194	0.021
3.1.5 The students' self-learning ability has been improved	4.73 ± 0.46	0.126	0.120	0.013
3.1.6 The students' independent thinking ability has been improved	4.73 ± 0.59	0.126	0.120	0.013
3.1.7 The students' ability to analyze and solve practical problems has been improved	4.73 ± 0.59	0.137	0.120	0.013
3.1.8 The students' critical thinking ability has been improved	4.60 ± 0.63	0.141	0.089	0.010
3.1.9 The students' cooperation ability has been improved	4.53 ± 0.64	0.110	0.061	0.007

4. Discussion

4.1. Reliability and scientific analysis

The Medical MOOC construction quality evaluation indicator system developed in this study is both scientifically sound and reliable. Firstly, it is based on the well-established theoretical framework of the Structure-Process-Outcome quality structure. Empirical evidence attests to the maturity of this theory in formulating evaluation indicators for medical teaching quality [15,18]. Additionally, the quality evaluation indicator system is meticulously crafted through a comprehensive literature review and subsequent refinement via two rounds of the Delphi method. Secondly, the credibility of the research findings is supported by the expertise and representativeness of the participating experts [16]. In this study, 15 experts were included from distinct regions across China, each recognized for their outstanding contributions and influence in medical education and MOOC construction. These experts provided invaluable insights and recommendations, demonstrating a remarkable degree of consensus and authoritative perspectives, and the outcomes of this research can be considered highly dependable [16].

4.2. Systematic and weight analysis

The medical MOOC construction quality evaluation indicator system developed in this study includes a comprehensive and systematic range of criteria. This approach aligns with Yang's perspective, emphasizing that e-learning evaluation should extend beyond assessing students' learning to include the evaluation of the e-learning environment, which encompasses the learning support platform, as well as the evaluation of learning resources [10]. A thorough analysis of these indicators, based on the "Structure-Process-Outcome" three-dimensional quality model, is provided below.

The structure indicator serves as the foundation of the course and is essential for the effective progression of medical MOOC teaching. This indicator comprises two components: platform technique and curriculum management. Our study results indicate that "curriculum management" (weighted at 0.750) carries greater significance than "platform technique" (weighted at 0.250). While "platform technique" does contribute to MOOC quality by influencing aspects such as interface design and personalized support, which directly impact students' learning experiences, "curriculum management" encompasses critical functions such as development and testing, organization and operation, and feedback and improvement [19]. Exceptional management of MOOCs is pivotal in attracting students and enhancing MOOC completion rates [20]. Therefore, "curriculum management" has been assigned a higher weight due to its greater impact on overall effectiveness. To ensure the smooth progression of Medical MOOC teaching, it is imperative to implement standardized and scientifically driven curriculum management, aligning with the objective of nurturing talent

[10,21]. Among the third-level indicators, “MOOC construction and administration are aided by adequate funding and personnel (0.030)” holds the highest combination weight, consistent with findings by King [22]. Unlike traditional programs and courses, MOOCs heavily rely on information technology and online platforms, involving not only instructors but also video producers, web designers, and web admins, among others. Therefore, sufficient human and financial support serves as a crucial foundation for structural factors. Moreover, traditional courses often entail more face-to-face interaction between teachers and students, enabling instructors to gather feedback through performance assessment and post-class communication. In contrast, MOOCs feature limited direct interaction between instructors and students, emphasizing the importance of timely feedback collection and adjustments in course quality evaluation. Consequently, “the feedback from students can be used to review or improve the course (0.026)” and “the course is periodically evaluated and revised to reflect the most recent developments in medical knowledge (0.019)” also carry significant combination weights.

The process indicator refers to the course construction and teaching implementation processes of medical MOOCs, covering teaching staff, course content, and student participation, which offer valuable insights into educational quality. Among the second-level indicators, “Course content (0.540)” had the highest weight and included course and teacher introductions, course outlines, learning objectives, and assessment methods [21]. All these elements collectively constitute the most critical components of MOOC content. In contrast, “student participation (0.163)” had the lowest weight. While learner engagement can somewhat reflect MOOC quality by increasing user numbers and fostering discussions, it is essential to acknowledge that learners’ commitment levels may vary. Some learners may engage selectively, focusing on specific chapters or initial course content, engaging in partial learning without intending to complete the entire medical MOOC [21]. Therefore, compared to “course content” and “teaching staff”, the quality of medical MOOC construction should not overly depend on students’ performance in the course. Among the third-level indicators, both “The assessment method and credit certification are clearly stated in the course” and “The course content can assist students in enhancing their self-learning ability” hold the highest combination weight of 0.054 each. Course assessment serves as a crucial tool for measuring learning outcomes, evaluating the achievement of teaching objectives, and assessing teaching quality. The results of assessments play a pivotal role in guiding teachers in designing effective teaching strategies [23]. Simultaneously, granting credits for MOOC scores provides positive incentives for students, fostering their motivation to learn [24,25]. Credit recognition is also a significant factor for students when selecting courses. Furthermore, MOOCs necessitate students to develop strong self-regulated learning abilities. Therefore, it is essential to adopt a student-centered and content-oriented approach that encourages active and continuous learning among students.

In the context of this study, the outcome indicator reflects the impact of curriculum construction, specifically through the second-level indicator “student achievement”. Within the third-level indicators, both “Students master the knowledge of the course” and “The students’ clinical thinking ability has been improved” hold the highest combined weight, both at 0.021. Clinical thinking ability pertains to a doctor’s capacity to make sound decisions based on real patient scenarios during patient care [26]. Developing the clinical thinking abilities of medical students within the classroom setting is crucial for their future clinical work. Among the third-level indicators in this category, “course completion rate” has the lowest combined weight at 0.004, followed by “the final grade of the course” at 0.007. Unlike traditional teaching methods, MOOC learners often possess diverse backgrounds and characteristics. One study suggests that when evaluating MOOC quality, the non-compulsory nature of MOOCs should be considered [21]. In this context, the quality of a MOOC is closely tied to individual student goals. Therefore, a MOOC can be considered of high quality if students have achieved their desired learning outcomes, even if they did not complete the entire course. Consequently, traditional metrics like completion rates may not be the most suitable quality indicators [7,20]. Furthermore, examination questions that assess overall achievement in medical MOOCs are typically objective, which can limit the assessment of students' subjectivity, initiative, and creativity. Consequently, it becomes more important to evaluate whether specific skills and learning abilities of students have been effectively enhanced rather than solely relying on course completion rates and overall achievement. This approach aligns with recent trends in Chinese medical education, which emphasize a shift towards "focusing more on competency and quality development" rather than a sole focus on knowledge transfer [27].

4.3. The effectiveness of combining the Delphi and AHP methods

In this present research study, a combination of the Delphi and AHP methods were applied to establish the evaluation index system for assessing the quality of Medical MOOC construction and to validate the weight coefficients assigned to each indicator. The Delphi technique, a well-established method for gathering diverse expert opinions and achieving consensus in a particular field of discussion, was utilized to develop the initial index system. The results obtained through the generalized Delphi method in this study were found to be highly accurate and reliable. In addition to the Delphi method, we incorporated the AHP, which quantifies experts' subjective judgments using mathematical models, facilitating a structured and scientific approach to multi-objective decision-making analysis [17]. The use of Saaty's scale for assigning significance greatly simplifies the process of expert judgment. Since our quality index system encompasses all facets and stages of medical MOOC teaching, the experts used a pairwise comparison matrix to determine the relative importance of each indicator. The results demonstrated that the consistency of all the indexes was confirmed through a consistency test, with a CR less than 0.10, indicating that the logical consistency was well-maintained, ensuring both scientific rigor and reliability. The quality index system for evaluating the construction of Medical MOOCs, developed through this combination of the Delphi and AHP methods, reflects the expertise of specialists and the precision of quantified assessment. A study that explored the efficacy of combining Delphi and AHP methods in developing evaluation systems found that this approach enhanced the reliability and validity of expert assessments, clarifying the qualifications required [28]. In summary, the quality index system we have established for evaluating the construction of Medical MOOCs, along with the assigned weight coefficients, can be considered scientifically sound, reliable, and practical.

5. Limitations

This study had several limitations that should be considered. First, due to resource and time constraints, we were only able to invite 15 experts to participate in the Delphi survey. While there is no consensus on the optimal panel size for a Delphi study, a systematic review by Paré [29]. suggested a typical range of 7 to 30 participants. Although our study falls within this range regarding panel size per round, it is important to note that we experienced a decline in the expert response rate, particularly in the first round, where the effective response rate was 73.33% (n = 11), which is slightly higher than the recommended threshold of 70% [30]. It is worth acknowledging that a drop in response rates is a common challenge encountered in e-Delphi studies [31].

Assessing the potential impact of attrition bias on the quality and credibility of our findings is a complex task. As the number of participants diminishes in subsequent rounds, there is a risk of overestimating the degree of consensus. Another limitation relates to the potential bias in our selection of experts. Most of the participating experts were from regions in China recognized for their robust evidence-based education, and this regional bias may affect the generalizability of our results. Furthermore, it is essential to emphasize that the indicator system developed in this study has not undergone validation, reliability testing, or empirical application. Therefore, this study recommends that future research include comprehensive quantitative evaluations of all indicators to enhance the effectiveness of medical MOOC teaching. We also propose conducting empirical studies on selected medical MOOCs using comprehensive fuzzy evaluation methods to further analyze and improve the applicability of this indicator system [30].

6. Conclusion

This study adopted a three-dimensional quality evaluation model as its theoretical framework and leveraged the Delphi method, drawing extensively from existing literature, to construct an evaluation indicator system for assessing the quality of medical MOOC construction, which ensured the scientific validity and credibility of the constructed system. The resultant evaluation indicator system comprises three primary indices, six secondary indices, and 71 tertiary indices. Furthermore, the study benefited from the expertise, motivation, and cooperation of the participating experts, reinforcing the scientific foundation of the medical MOOC construction quality evaluation index system. The AHP was employed to determine the weighted values across all levels of the quality index, which were established based on experts' assessments using Saaty's scale, and the consistency ratio was found to be less than 0.10, indicating the reliability of the weight coefficients. The analysis of the weighted values within the index system revealed that both course content and teaching staff significantly influence the quality of medical MOOC construction. Additionally, effective curriculum management practices contribute to the enhancement of medical MOOC teaching quality. Specifically, the combination of well-designed course content and efficient teacher management serves as the cornerstone for achieving high-quality medical MOOC teaching. In summary, the proposed indicator system in this present study provides a valuable and practical tool for continuously enhancing the quality of medical MOOCs in China.

Funding

Xizang Education Science Research Project 2023 (Project No.: XZEDGP230066); Nursing Discipline Research Project of the Chinese Medical Association Publishing House for 2022–2023 (Project No.: CMAPH-NRD2022012)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Williams R, 2024, An Overview of MOOCs and Blended Learning: Integrating MOOC Technologies Into Traditional Classes. IETE Journal of Education, 65(2): 84–91.

[2] Ahmad I, Sharma S, Singh R, et al., 2022, MOOC 5.0: A Roadmap to the Future of Learning. Sustainability, 14(18): 11199.

[3] Pichayakul P, 2024, Academic Outreach During COVID-19 Lockdown at a University Level: Using MOOC as a Strategy. Journal of Education and Innovation, 26(3): 86–99.

[4] Guo Y, 2014, MOOCs and the Future of China's Higher Education. Journal of Higher Education Management, 5: 29–33.

[5] People's Daily, 2023, Technology-Enabled Education Shares University Resources, visited on 2026/6/11,http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2023/2023_zt01/mtbd/202302/t20230212_1043924.html.

[6] Zhu H, Xu J, Wang P, et al., 2023, The Irreplaceable Role of Medical Massive Open Online Courses in China During the COVID-19 Pandemic. BMC Medical Education, 23(1): 323.

[7] Huang H, Jew L, Qi D, 2023, Take a MOOC and Then Drop: A Systematic Review of MOOC Engagement Pattern and Dropout Factor. Heliyon, 9(4): e15220.

[8] Celik B, Cagiltay K, 2024, Uncovering MOOC Completion: A Comparative Study of Completion Rates From Different Perspectives. Open Praxis, 16(3): 445–456.

[9] Zhou L, Tang M, Liu J, 2025, Analysis of Factors Influencing MOOC Quality Based on I-DEMATEL-ISM Method. Systems and Soft Computing, 7: 200220.

[10] Yang L, Zou J, Gao J, et al., 2023, Assessing the Effectiveness of Massive Open Online Courses on Improving Clinical Skills in Medical Education in China: A Meta-Analysis. Heliyon, 9(8): e19263.

[11] Cheng J, Yuen K, Chiu W, 2023, Systematic Review of MOOC Research in Mainland China. Library Hi Tech, 41(5): 1476–1497.

[12] Jain R, 2023, Premchand's The Chess Players: A Comparative Analysis of YouTube and MOOC Platforms. New Literaria, 4(1): 45–52.

[13] Ponnaiah M, Bhatnagar T, Ganeshkumar P, et al., 2022, Design and Implementation Challenges of Massive Open Online Course on Research Methods for Indian Medical Postgraduates and Teachers – Descriptive Analysis of Inaugural Cycle. BMC Medical Education, 22(1): 369.

[14] Shang Z, 2023, Use of Delphi in Health Sciences Research: A Narrative Review. Medicine, 102(7): e32829.

[15] Robby A, Nurjazuli N, Izzah L, et al., 2025, From Structure to Outcome: The Role of the Donabedian Model in Global Health Service Management Research (1987–2024). E3S Web of Conferences, 650: 02025.

[16] Hsu C, Sandford A, 2007, The Delphi Technique: Making Sense of Consensus. Practical Assessment, Research and Evaluation, 12(1): 10.

[17] Saaty L, 1990, How to Make a Decision: The Analytic Hierarchy Process. European Journal of Operational Research, 48(1): 9–26.

[18] Kelly L, Whitaker A, Burchinal M, et al., 2026, Testing the Structure–Process–Outcome Model Across Different Early Care and Education Policy Contexts. Developmental Psychology, 62(2): 442–461.

[19] Chen H, 2023, Influencing Factors of the Quality of MOOCs Based on the KANO Model. International Journal of Emerging Technologies in Learning, 18(17): 20–32.

[20] Mehrabi M, Safarpour R, Keshtkar A, 2022, Massive Open Online Courses (MOOCs) Dropout Rate in the World: A Protocol for Systematic Review and Meta-Analysis. Interdisciplinary Journal of Virtual Learning in Medical

Sciences, 13(2): 85–92.

[21] Hendriks A, de Jong M, Admiraal F, et al., 2020, Instructional Design Quality in Medical Massive Open Online Courses for Integration Into Campus Education. Medical Teacher, 42(2): 156–163.

[22] King I, Lee I, 2023, MOOC Policies Across the Globe. In: King I, Lee W I (eds.) A Decade of MOOCs and Beyond. Springer International Publishing: 61–88.

[23] Liu X, Feng J, Liu C, et al., 2023, Medical Education Systems in China: Development, Status, and Evaluation. Academic Medicine, 98(1): 43–49.

[24] Moore R, Blackmon S, 2022, From the Learner's Perspective: A Systematic Review of MOOC Learner Experiences (2008–2021). Computers & Education, 190: 104596.

[25] Oo S, Schofield S, Aung T, et al., 2022, Integrating a Mobile-Learning Platform for Enhancing Clinical Teaching: The Learners' Perspective. International Journal of Learning, Teaching and Educational Research, 21(11): 87–111.

[26] Hui Z, Zewu Z, Jiao H, et al., 2025, Application of ChatGPT-Assisted Problem-Based Learning Teaching Method in Clinical Medical Education. BMC Medical Education, 25(1): 50.

[27] Sun H, Chen H, Dong C, et al., 2021, Building a First-Class Nursing Major With National Standards for Nursing Teaching Quality. Chinese Journal of Nursing Education, 18(5): 389–394.

[28] Meng Q, Wang F, Chen K, et al., 2026, Development and Evaluation of an Evidence-Based Management Indicator System for Transitional Care in Preterm Infants Using a Delphi and AHP Approach. Journal of Nursing Management, 2026(1): 2638144.

[29] Paré G, Cameron F, Poba-Nzaou P, et al., 2013, A Systematic Assessment of Rigor in Information Systems Ranking-Type Delphi Studies. Information & Management, 50(5): 207–217.

[30] Hayes B, Fitzgerald D, Doherty S, et al., 2015, Quality Care, Public Perception and Quick-Fix Service Management: A Delphi Study on Stressors of Hospital Doctors in Ireland. BMJ Open, 5(12): e009564.

[31] Donohoe H, Stellefson M, Tennant B, 2012, Advantages and Limitations of the e-Delphi Technique: Implications for Health Education Researchers. American Journal of Health Education, 43(1): 38–46.