

Developing Evaluation Criteria for Competency-Based Curriculum in Medical Colleges

So Young Lee

Chungnam National University College of Medicine

Seung-Hee Lee (✉ lshcho@snu.ac.kr)

Seoul National University College of Medicine

Jwa-Seop Shin

Seoul National University College of Medicine

Jooah Kim

Yonsei University College of Dentistry

Hyun Bae Yoon

Seoul National University College of Medicine

Research Article

Keywords: Curriculum evaluation, Evaluation criteria, Competency-based curriculum, Delphi, CIPP evaluation model, Kirkpatrick evaluation model, Continuous Quality Improvement

Posted Date: November 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1045926/v1>

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

[Read Full License](#)

Developing Evaluation Criteria for Competency-Based Curriculum in Medical Colleges

So Young Lee¹, Seung-Hee Lee^{2,3*}, Jwa-Seop Shin^{2,3}, Jooah Kim⁴, and Hyun Bae Yoon^{3,5}

1. Department of medical education Chungnam National University College of Medicine, Daejeon, Korea
2. Department of Medical Education Seoul National University College of Medicine, Seoul, Korea
3. Department of Human Systems Medicine Seoul National University College of Medicine, Seoul, Korea
4. Department of Dental Education in Yonsei University College of Dentistry, Seoul, Korea
5. Office of Medical Education Seoul National University College of Medicine, Seoul, Korea

* Correspondence: lshcho@snu.ac.kr

Abstract

Background

The present study was conducted to examine the operation status of the Competency-Based Curriculum (CBC), which has become the main curriculum in the medical schools in Korea, and prepare valid evaluation criteria consented by experts to make improvements. The evaluation criteria were derived based on a model built by combining the Context, Input, Process and Product (CIPP) evaluation model and the Kirkpatrick evaluation model, which are representative educational evaluation models.

Methods

Firstly, literature survey was performed and a semi-structured interview was conducted with 5 experts to develop a draft of the evaluation criteria. To verify the validity of the draft of the developed evaluation criteria, two surveys based on the modified Delphi methodology were conducted with a panel consisting of 20 experts.

Results

Based on the literature survey and the expert interview, a draft of the evaluation criteria was derived, including 5 evaluation areas, 18 evaluation items and 58 evaluation indicators. Two Delphi surveys were conducted to validate the evaluation criteria. The evaluation criteria that showed a relatively low content validity ratio (CVR) were corrected and complemented by reflecting the experts' opinions to finally derive 5 evaluation areas, 16 evaluation items and 51 evaluation indicators.

Conclusions

The significance of the present study is that an evaluation model and its evaluation criteria suitable for the curriculums of the medical schools in Korea were developed with the consensus of the experts. The preparation of valid evaluation criteria will open up the possibility of improving the evaluation in medical schools, contributing to the improvement of the educational quality and the continued quality improvement

of medical education.

Keywords

Curriculum evaluation, Evaluation criteria, Competency-based curriculum, Delphi, CIPP evaluation model, Kirkpatrick evaluation model, Continuous Quality Improvement

Background

To improve the quality of education, many medical schools have newly introduced the Competency-Based Curriculum (CBC), which has become the center of the current curriculums of medical schools (1) (2) (3). The CBC is a curriculum model that defines the competencies that should be exercised at the work sites after the learners finish their learning and that plans the educational goals, contents and methods, and evaluation accordingly (4) (5) (6) (7). The CBC started to be recognized in Korea in the middle and late 2000s (8). Now that a decade has passed since the CBC was introduced and settled in the medical schools in Korea, it is necessary to examine if the education is implemented properly to make needed improvements.

Previous studies reveal the attempts to assess and improve the medical education in Korea (9) (10) (11) (12). However, most of the previous studies were conducted without a systematic framework, and the CBC has not been sufficiently investigated.

The evaluation criteria are so important that they can determine the success of evaluation (13). Therefore, valid evaluation criteria must be prepared to improve the curriculum. Nevertheless, the field of medical education in Korea lacks the evaluation criteria prepared through a consensus.

Various evaluation models for education evaluation have been suggested. Many scholars, including Worthen et al. (1997) (14), Russ-Eft & Preskill (2009) (15) and Cook (2010) (16), presented various viewpoints to classify the models. However, despite the various ways of classification, the models of which usefulness have been proved by evaluation in various education fields, including medical education, are the Context, Input, Process and Product (CIPP) evaluation model and the Kirkpatrick evaluation model

(17).

The CIPP evaluation model is for the educational evaluation in the aspects of context, input, process and product to provide information needed for decisions. The global society of medical education has introduced the CIPP evaluation model, which is widely applied to educational evaluation(18) (19) (20) (21) (22) (23) (24). However, the previous studies regarding the CIPP evaluation model were not conducted by reflecting the CBC of medical schools.

The 4-level Kirkpatrick evaluation model is one of the popular educational evaluation models that are applied to the evaluation of the results of medical education (25) (26) (27). This model reviews the results of education in 4 levels, which are reaction, learning, behavior and result (28). However, the studies based on this evaluation model have a limitation that the positive and negative details that happen throughout the educational process may not be examined (29) (30) (31).

The CBC has been developed based on the behaviorism theory which is interested in the change of the learner's behavior (32) (33). Behaviorism is focused on the learner's observable behavior (34), which is consistent with the idea presented by Kirkpatrick (2009) that no behavior change occurs without learning. Therefore, the Kirkpatrick evaluation model can provide useful information to the CBC (35).

Studies have shown that the combination of the CIPP evaluation model and the Kirkpatrick evaluation model gives a complementary synergic effect (36) (37). Kim et al. (2011) combined the two models in pursuit of comprehensive understanding of the operational process of education and the results of education (36). Kim (2018) sought both the outcome and process but eventually focused on the outcome-oriented Kirkpatrick evaluation model (37). When the two models, which are particularly useful in the evaluation of medical education, are combined and applied to the development of evaluation criteria, the quality of medical education is expected to improve.

In summary, the improvement of the CBC of the medical schools requires the development of comprehensive and systematic evaluation criteria for the verification of outcome in the entire educational process. The CIPP evaluation model and the Kirkpatrick evaluation model need to be applied to support the comprehensive and systematic evaluation.

Methods

The CBC evaluation criteria for medical schools were developed in the following stages: the drafting of evaluation criteria; and the validation of evaluation criteria by Delphi method. The states are described in details below.

Drafting of Evaluation Criteria

The evaluation criteria drafting stage is to prepare the draft of the evaluation criteria. For this, previous studies were analyzed with regard to the concept, purpose and characteristics of the CBC and the evaluation criteria for the evaluation of medical education programs, and a semi-structured interview was conducted with 5 experts in medical education and education evaluation. Table 1 shows the experts with whom the interview was conducted. All the experts were current professors who had a career of at least 14 years in the field of medical education and experiences specialized in the introduction, operation and evaluation of the CBC. Their majors were medicine and pedagogy. The interview was conducted between July 29 and August 10 in 2020 with one expert each time for about 1 hour to 1 and half hour. The materials acquired from the interview were analyzed and interpreted by the comprehensive analytical procedures utilized by Koo (2018) and provided by Lee and Kim (2014) (38, 39). The entire recording of the interview was dictated, and the collected materials were repeatedly read from the beginning to the end to list meaningful statements, establish units and prepare a list of them.

<Table 1> Experts with whom the interview was conducted.

Expert groups	Number of experts	Occupation	Total career in medical education
Experts in medical education	4	Professor	14 to 22 years
Experts in education evaluation	1	Professor	14 to 22 years

The evaluation criteria extracted from previous studies and the analytical results of the expert interview were consolidated and summarized to derive the draft of the evaluation criteria. The draft was modified by 3 experts who majored in medical education, pedagogy and medicine to establish the draft of the evaluation criteria.

Validation of Evaluation Criteria

In the present study, after the draft was prepared, the evaluation criteria were validated by applying the modified Delphi method.

Two Delphi surveys were conducted to verify the validity of the drafted evaluation criteria. Because the Delphi expert panel should include 10 to 15 experts to minimize the errors and secure the reliability (40) (41) (42), a Delphi panel was constituted in the present study by including 20 medical education experts who majored medical education, pedagogy and medicine. Table 2 shows the details about the Delphi panel. The Delphi panel was formed by selecting the experts who were basically in the medical education at present and had at least a doctor's degree (course completion) and experiences of operating and evaluation the CBC at actual medical education sites for 3 years or longer. The Delphi panel consisted of experts who majored in medical education, those who majored in both medicine and pedagogy, those who majored in pedagogy, and those who majored in other fields. Four of the experts had a career in the medical education for 5 years or less, 8 between 6 and 10 years, and 8 for 11 years or longer. The questionnaire was received via e-mail, and all the 20 members of the Delphi panel (100%) participated in the first and second surveys.

<Table 2> Inclusion criteria and number of experts in the panel group.

Panel group	Total career in medical education	Number of selected experts	Number of experts who responded in the Delphi surveys	
			First	Second
Majored in medical education	10 to 11 years	2	2	2
Majored in both medicine and pedagogy	13 years	1	1	1
Majored in pedagogy	3 to 20 years	10	10	10
Majored in medicine	4.5 to 20 years	6	6	6
Majored in other fields	10 years	1	1	1

Two Delphi surveys were implemented for the validation of the evaluation criteria. The expert panel was asked to respond in a 5-point Likert scale (1-Very invalid, 2-invalid, 3-Moderate, 4-Valid, and 5-Very valid) about the validity of the evaluation areas, evaluation items and evaluation indicators in the draft evaluation criteria derived in the evaluation criteria drafting stage. The judgment of the validity was based on the review of the basic statistics including the means, standard deviations and quartiles. In particular, the judgment was made according to the content validity ratio (CVR), which is an essential variable in the validity judgment in Delphi surveys, by deciding the CVR is higher than 0.42 as suggested by Lawshe (1975) for an expert panel consisting of 20 members (43).

In addition, since the validity in the Delphi method may be presented by analyzing the levels of opinion convergence and consensus by the expert panel (44) (45) (46), a degree of consensus of 0.75 or higher and a degree of opinion convergence of 0.5 or lower were considered as a high validity.

In the Delphi method, the stopping criterion that determines the rounds was the Coefficient of Variation (CV), which was considered to examine the stability. The CV is the standard deviation divided by the arithmetic mean. A CV of 0.5 or lower was considered as requiring no additional round, 0.5 to 0.8 was

considered as a relatively stable, and 0.8 or higher as requiring an additional survey (47) (48).

The data from the first and second Delphi survey were analyzed by calculating the frequency, percentage, mean, standard deviation, median, quartile, degree of consensus, degree of convergence and CV by using Excel 2016 software program.

In addition, the Delphi panel was asked to freely describe the parts of the individual questions that required correction, addition or removal. After completing a round, the opinions from the Delphi panel were summarized to modify the questions, especially those that were pointed out by at least two experts. The questions that were pointed out by at least one expert were modified through the consultation with the experts who reviewed the draft evaluation criteria.

Results

Development of Draft Evaluation Criteria

A literature survey and an expert interview were conducted to prepare the draft of the evaluation criteria. The evaluation criteria of the medical education programs were explored through the literature survey. The focus of the CVC evaluation of the medical school in Korea was summarized through the expert interview to prepare the evaluation areas, evaluation items and evaluation indicators. Sixteen previous reports were found from the literature survey to be reviewed in the present study in relation to the evaluation criteria for education programs (49) (50) (51) (52) (53) (54) (18) (55) (56) (29) (20) (21) (57) (23) (58) (59). In addition, although not included in the field of medical education, previous reports from other fields were also reviewed since they clearly present the evaluation criteria of the Kirkpatrick evaluation model (60) (61).

Table 3 summarizes the results of the literature survey and the expert interview. According to the literature survey and the expert interview, the context, input and process of the CIPP evaluation model were applied as evaluation area without modification. The product area was divided into the learning outcome and curriculum quality management, and the Kirkpatrick evaluation model was applied to the learning outcome part, building a mixed model and deriving the evaluation areas for the model. The evaluation items and

evaluation indicators were also prepared as the first Delphi evaluation criteria shown in Table 5 and Appendix 1 based on the literature survey and the expert interview.

<Table 3> the mixed evaluation model and evaluation areas derived from the literature survey and the expert interview.

Evaluation model	Evaluation area		Evaluation content	
<i>CIPP evaluation model</i>	Context		The educational goals and competences are determined, and the demands of the medical school stakeholders, obstacles (problems), and the conditions and environment around the education are evaluated.	
	Input		The appropriateness of the curriculum/instruction design and the distribution of the human and material resources to implement the same are evaluated.	
	Process		The details and levels of practice relative to the goal-based plan are evaluated, and the system for the practice evaluation and the evaluation implementation are evaluated.	
	Product	Evaluation of learn outcome <i>(Kirkpatrick evaluation model)</i>	reaction	The change is verified in the knowledge, skills and attitude that learners acquire by participating in the education, and
		learning	The change is evaluated in the knowledge, skills and attitude that learners acquire by participating in the education.	

Evaluation model	Evaluation area			Evaluation content	
			behavior	the effect on the actual practice through the application after the learning is evaluated.	The application and transfer of the knowledge, skills and attitude that learners acquire through the education are evaluated. In addition, the support of the learners' application by the environmental factors is evaluated.
			result		The degree of substantial contribution by the learners to the organization through the education and the degree of the organizational improvement are evaluated.
		Curriculum quality management	product	The sustainability of the success factors of the educational program and the points requiring improvement are verified. The procedures and systems for the improvement of the curriculum are evaluated.	

Validation of Evaluation Criteria (Delphi)

Results of First Delphi Survey

The five evaluation areas of the evaluation criteria were 1. context, 2. Input, 3. Process, 4. Learning outcome and 5. Outcome of curriculum. In the first Delphi survey, as shown in Table 4, the mean value was 4.0 points or higher in all the evaluation areas and the standard deviation was 0.5 to 0.74. The median and the mode were both 5 points, indicating that the validity of the evaluation areas was very high. The overall rate of positive response was 95% to 100%. The CVR was 0.90 to 1.00, which was higher than the validity level (0.42) suggested by Lawshe (1975) for an expert panel consisting of 20 members (43). Therefore, the survey showed that all the evaluation areas presented in the draft evaluation criteria were valid.

The comments of the expert panel on the individual evaluation areas were additionally reviewed, and 1. Context was modified to 1. Educational environment & context, and 5. Outcome of curriculum to 5. Continuous Quality Improvement(CQI) according to their comments.

The mean score of most evaluation items (17 items) was 4 points or higher, except the evaluation item of '1.2 Design principles.' The mode was 5 points except two items, and the mean was 4 points or higher in 17 items except '1.2 Design principles.'

The only evaluation item that failed to exceed the CVR value of 0.42, the validity criterion, was '1.2 Design principles.' To treat the evaluation item, the name and the place of the criterion were adjusted by summarizing the opinions stated by the Delphi panel. In other words, '1.2 Design principles' was moved to 2.1. On the other hand, the evaluation items of which validity was secured as the CVR was over 0.42 but showed a relatively low CVR value were also modified by reflecting the opinions from the expert panel. '4.3 Application of acquired competence' was changed to 'Performance transfer of learning.' '1.4 Organization' was changed to 'Administrative organization.' '2.1 Educational contents and educational methods' was divided into separate items, which are 'Educational contents' and 'Educational methods.' '2.2 Educational strategy' was removed, and the lower level evaluation indicator was divided into 'Educational contents' and 'Educational methods.' '4.4 Performance accomplishments by graduates' was changed to 'Performance accomplishments by graduates after graduation.'

Appendix 1 shows the results of the analysis of the evaluation indicators derived from the first Delphi survey. Among the total of 58 evaluation indicators, 47 indicators showed a mean value of 4.0 or higher, and the other 11 indicators showed a mean value between 3.0 and 4.0, mostly over 3.5, indicating that the overall validity of the evaluation indicators was high. The mode and the median were between 4 and 5, except 3 indicators. The CVR of 53 indicators was over 0.42, except 5 indicators, and thus the validity was secured.

Most of the evaluation indicators included in the area of '1. Context' showed a mean value of 4.0 or higher, and their CVR value also satisfied the validity criterion. However, '1.4.1 Sufficiency of the will and leadership of executives' showed a CVR value of 0.40, which failed to satisfy the validity criterion.

Most of the evaluation indicators included in the area of '2. Input' showed a mean value of 4.0 or higher, and their CVR value also satisfied the validity criterion. However, '2.3.6 Appropriateness of introducing absolute evaluation' showed a mean value of 3.58 and a CVR value of 0.06, which failed to satisfy the validity criterion.

Most of the evaluation indicators included in the area of '3. Process' showed a mean value of 4.0 or higher, and their CVR value was distributed between 0.60 and 0.80. However, '3.1.1. Appropriateness of competence acquisition opportunity in clinical practice' in the evaluation item of '3.1 Teaching-learning process' showed a mean value of 4.0 or lower and a CVR value of 0.16, which failed to satisfy the validity criterion.

In the area of '4. Learning outcome,' 8 out of the 10 evaluation indicators showed a mean value between 3 and 4, and the CVR value satisfied the validity criterion. On the other hand, '4.4.2 Change in patients and treatment' and '4.4.3 Change in society and science' failed to satisfy the validity criterion.

In the area of '5. Curriculum,' all the indicators of '5.1 Sustainability and feedback' showed a mean value of 4.0 or higher and a CVR value of 0.7 or higher.

Summarizing the opinions of the expert panel on the evaluation indicators obtained through the first Delphi survey, the indicators showing a CVR of 0.42 or lower were removed, and the removed indicators were '1.4.1 Sufficiency of the will and leadership of executives,' '2.3.6 Appropriateness of introducing absolute evaluation,' '3.1.1. Appropriateness of competence acquisition opportunity in clinical practice,' '4.4.2 Change in patients and treatment' and '4.4.3 Change in society and science.'

In addition, according to the opinions of the Delphi expert panel on the key modifications, the terms were firstly corrected. A representative example is '1.1.2 Appropriateness of collecting stakeholders' demands in educational competences and education goals.' Second, the evaluation indicators were moved to appropriate positions. Specifically, as '2.1 Educational contents and educational methods' was divided into separate items, which are 'Educational contents' and 'Educational methods' and '2.2 Educational strategy' was removed, the related educational indicators were moved. The indicators of '2.2 Educational strategy' were allocated to the newly created items of '2.1 Educational contents' and '2.2 Educational methods.' Third, the concepts that had to be added were reflected to the evaluation items. '2.2.4 Appropriateness of

basic medicine-clinical medicine education connection' was modified to 'Appropriateness of basic medicine- medical humanities-clinical medicine education connection' according to the opinion that the connection with the medical humanities should also be considered. In addition, '2.3.7 Appropriateness of remediation provision' was modified to '2.3.6. Appropriateness of providing remediation and reevaluation' according to the opinion that the reevaluation should also be considered. Fourth, the redundant parts of the evaluation indicators were integrated into a single indicator. '3.3.3 Sustainability of the activities by the competence review committee' and '3.3.4 Appropriateness of curriculum supervising professor' were integrated into '3.3.3 Sustainability of the activities by the competence review committee, the curriculum development committee, and the education evaluation committee.'

Results of Second Delphi Survey

The second Delphi survey was conducted to verify the validity of the evaluation area, evaluation items and evaluation indicators of the evaluation criteria modified according to the analysis of the results from the first Delphi survey. Table 4 shows the results of the second Delphi survey.

The mean validity of all the evaluation areas was 4.0 or higher in the 5-point scale. The median and the mode were both 5 points, indicating the validity of the evaluation areas was very high. The overall rate of positive response was 95% to 100%. The CVR was 0.90 to 1.00, which was higher than the validity level (0.42). Therefore, all the presented evaluation areas were found valid.

In addition, the degree of consensus of all the evaluation areas was over the criterion (0.75), and their degree of convergence was 0.5 or lower, satisfying the validity conditions. The CV value was 0.5 or lower, indicating that survey results could be considered as the final results without the need for conducting an additional round of Delphi survey.

Beside the analysis of the basic statistics, the additional comments of the expert panel on the individual evaluation areas were reviewed. As a result, '1. Educational environment & context' was recovered to '1. Context' according to the major opinion.

The mean score of 16 evaluation items was 4 or higher, except the two evaluation items of '2.1 Curriculum design principles' and '4.4 Performance accomplishments by graduates after graduation.' The mode and the media of 15 evaluation items were 5 points, except the one item of '4.3 Performance transfer of

learning,' indicating that the validity was very high. One evaluation item that failed to exceed the CVR criterion of 0.42 for an expert panel consisting of 20 member was '4.4 Performance accomplishments by graduates after graduation,' which was therefore removed.

All the evaluation items exceeded the consensus criterion of 0.75 and the convergence criterion of 0.5, proving their validity. The CV value was 0.5 or less, indicating that that survey results could be considered as the final results without the need for conducting an additional round of Delphi survey.

The names and positions of the evaluation criteria were changed by reflecting the opinions from the Delphi panel members. Specifically, since '3.3 Education monitoring status and monitoring system' and '5.1 Continuous curriculum quality assurance and improvement' are similar concepts, '3.3 Education monitoring status and monitoring system' was removed while retaining '5.1 Continuous curriculum quality assurance and improvement.' The names of the evaluation items were changed: '1.2 Teacher and learner analysis' to '1.2 Teacher and learner'; '2.4 Learning evaluation design' to '2.4 Learning evaluation'; '3.2 Achievement evaluation' to '3.2 Achievement evaluation process'; '4.2 Competence achievement' to '4.2 Degree of competence achievement'; and '4.3 Performance transfer of learning' to '4.3 Transfer of learning.'

As shown in Appendix 1, among the total of 52 evaluation indicators, 47 indicators showed a mean value of 4.0 or higher, and the other 5 indicators showed a mean value over 3.5, indicating that the overall validity of the evaluation indicators was high. The mode and the median were between 4 and 5. The CVR of all the indicators showed that the validity was secured.

All the evaluation indicated exceeded the consensus criterion of 0.75 and the convergence criterion of 0.5, proving their validity. The CV value was 0.5 or less, indicating that that survey results could be considered as the final results without the need for conducting an additional round of Delphi survey.

In addition, unnatural expressions were corrected according to the demands of the Delphi experts for the correction. Representative examples include '1.2.2 Appropriateness of analyzing stakeholders' demands in educational competences and education goals,' '1.1.4 Clarity of recognition of educational competences and education goals,' and '2.3.4 Sufficiency of providing education opportunities through meetings with real patients.'

Second, the evaluation indicators were moved to appropriate positions and their expressions were properly corrected. First, as the evaluation item of '3.3 Education monitoring status and monitoring system' was removed, '3.3.3 Sustainability of the activities by the competence review committee, the curriculum

development committee, and the education evaluation committee,' included therein, was removed, and a part of the details of '5.1.2 Appropriateness of curriculum evaluation method' was included instead. In addition, '3.3.1 Appropriateness of education evaluation method' and '3.3.2 Collection of education evaluation results' were moved to '5.1.2' and '5.1.4,' respectively. In addition, '2.5.2 System of organization in charge of medical education' and '2.5.3 System of organization in charge of evaluation' were respectively moved to '1.3.4' and '1.3.5' and the expressions were properly corrected according to the opinions from the Delphi panel.

Table 6 shows the finally evaluation criteria derived from the second Delphi survey, including the 5 area, 16 items and 51 indicators.

<Table 4> Results of first and second Delphi surveys on the evaluation areas.

First survey						Second survey					
Evaluation area	Mean	CV	Degree of consensus	Degree of convergence	CV	Evaluation area	Mean	CV	Degree of consensus	Degree of convergence	CV
1. Context	4.55	0.9	0.8	0.5	0.13	1. Educational environment & context	4.75	1	1	0	0.09
2. Input	4.55	1	0.8	0.5	0.11	2. Input	4.85	1	1	0	0.08
3. Process	4.6	0.9	0.8	0.5	0.13	3. Process	4.7	0.9	0.85	0.38	0.12
4. Learning outcome	4.55	0.9	0.8	0.5	0.13	4. Learning outcome	4.75	0.9	1	0	0.12
5. Outcome of curriculum	4.45	0.9	0.8	0.5	0.17	5. Continuous Quality Improvement	4.8	0.9	1	0	0.11

m						ent					
---	--	--	--	--	--	-----	--	--	--	--	--

<Table 5> Results of first and second Delphi surveys on the evaluation items.

First survey						Second survey					
Evaluation item	Mean	CV R	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CV R	Degree of consensus	Degree of convergence	CV
1.1 Competences and goals	4.6	1	0.80	0.50	0.11	1.1 Educational competences and educational goals	4.8	0.9	1.00	0.00	0.11
1.2 Design principles	3.8	<u>0.2</u>	0.69	0.63	0.24						
1.3 Teachers and learners	4.55	0.9	0.80	0.50	0.13	1.2 Analysis of teachers and learners	4.75	0.9	1.00	0.00	0.12
1.4 Organization	4.2	0.5	0.72	0.63	0.23	1.3 Administrative organization	4.5	0.8	0.80	0.50	0.15
						2.1 Curriculum design principles	3.85	0.7	1.00	0.00	0.21
2.1 Educational	4.35	0.5	0.75	0.63	0.23	2.2 Educational	4.8	1	1.00	0.00	0.09

First survey						Second survey					
Evaluation item	Mean	CV	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CV	Degree of consensus	Degree of convergence	CV
contents and educational methods						contents					
2.2 Educational strategy	4.25	0.6	0.80	0.50	0.26	2.3 Educational methods	4.8	1	1.00	0.00	0.09
2.3 Evaluation	4.3	0.7	0.78	0.50	0.20	2.4 Learner evaluation design	4.75	0.9	1.00	0.00	0.12
2.4 Human resources	4.5	0.8	0.80	0.50	0.15	2.5 Teacher support	4.8	1	1.00	0.00	0.09
2.5 Material resources and facilities	4.4	0.7	0.80	0.50	0.17	2.6 Learner support	4.7	0.9	1.00	0.00	0.12
2.6 Educational support	4.4	0.8	0.80	0.50	0.19	2.7 Material resources	4.4	0.8	0.80	0.50	0.23
3.1 Teaching-learning process'	4.4	0.8	0.80	0.50	0.21	3.1 Teaching-learning process'	4.7	0.9	1.00	0.00	0.12
3.2 Evaluation of learners	4.7	1	0.80	0.50	0.10	3.2 Achievement evaluation	4.6	0.7	1.00	0.00	0.19

First survey						Second survey					
Evaluation item	Mean	CV R	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CV R	Degree of consensus	Degree of convergence	CV
3.3 Education monitoring and improvement	4.45	0.8	0.80	0.50	0.19	3.3 Education monitoring status and monitoring system	4.65	0.8	1.00	0.00	0.14
4.1 Degree of satisfaction	4.35	0.7	0.80	0.50	0.20	4.1 Degree of satisfaction	4.45	0.9	0.80	0.50	0.21
4.2 Competence acquisition	4.4	0.8	0.80	0.50	0.19	4.2 Competence achievement	4.84	1	1.00	0.00	0.08
4.3 Application of acquired competences	4.1	0.5	0.69	0.63	0.22	4.3 Performance transfer of learning	4.15	0.6	0.75	0.50	0.21
4.4 Performance accomplishments by graduates	4	0.5	0.69	0.63	0.21	4.4 Performance accomplishments by graduates after graduation	3.9	<u>0.4</u>	0.50	1.00	0.26
5.1 Sustainability and	4.45	0.8	0.80	0.50	0.15	5.1 Continuous curriculum	4.75	0.9	1.00	0.00	0.12

First survey						Second survey					
Evaluation item	Mean	CV	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CV	Degree of consensus	Degree of convergence	CV
feedback						quality assurance and improvement					

<Table 6> Evaluation criteria finally derived by reflecting the results of the second Delphi survey.

Evaluation areas	Evaluation items	Evaluation indicators
1. Context	1.1 Competences and educational goals	1.1.1. Clarity of educational competences and educational goals
		1.1.2. Appropriateness of analyzing stakeholders' demands in educational competences and education goals
		1.1.3. Sufficiency of reflecting essential medical education competences to educational competences an education goals
		1.1.4. Clarity of recognition of educational competences and educational goals
	1.2 Teachers and learners	1.2.1. Integrity of learner analysis
		1.2.2. Integrity of teacher analysis
	1.3 Administration and organization	1.3.1 Integrity of finance and regulations for curriculum implementation
		1.3.2. Sufficiency of member consensus on curriculum introduction

Evaluation areas	Evaluation items	Evaluation indicators
		1.3.3. System of organization in charge of medical education
		1.3.4. System of organization in charge of evaluation
2. Input	2.1 Curriculum design principles	2.1.1. Sufficiency of reflecting curriculum design principles
		2.1.2 Scrupulosity of syllabus
		2.1.3 Appropriateness of amount of learning goals
	2.2 Educational contents	2.2.1. Appropriateness of educational contents
		2.2.2 Appropriateness of basic medicine-medical humanities-clinical medicine education connection
		2.2.3 Appropriateness of balance coordination in curriculum and subjects
	2.3 Educational methods	2.3.1. Appropriateness of education methods
		2.3.2 Appropriateness of educational programming in consideration of clinical performance capabilities
		2.3.3. Appropriateness of clinical educational programming in consideration of learner's level and education after graduation
		2.3.4. Sufficiency of providing education opportunities through meetings with patients
		2.3.5. Practice reinforcement
	2.4 Learner evaluation	2.4.1. Clarity of establishing basic principles of learner evaluation
		2.4.2. Appropriateness of learner evaluation methods
2.4.3. Evaluation focused on performance		
2.4.4. Clarity of competence achievement evaluation criteria		

Evaluation areas	Evaluation items	Evaluation indicators
		2.4.5. Appropriateness of implementing progress test
		2.4.6. Appropriateness of providing remediation and reevaluation
		2.4.7. Appropriateness of portfolio evaluation
		2.4.8. Appropriateness of formative evaluation and summative evaluation
	2.5 Teacher support	2.5.1. Clarity of granting authorities and roles to curriculum supervising professor
		2.5.2. Appropriateness of support for teaching capability development
	2.6 Learner support	2.6.1. Appropriateness of learning support for learners
	2.7 Material resources	2.7.1. Sufficiency of facilities and equipment
		2.7.2. Organic connection between medical schools and hospitals for practical training
	3. Process	3.1 Teaching-learning process
3.1.2 Smoothness of conducting education		
3.1.3. Appropriateness of feedback		
3.2 Achievement evaluation process		3.2.1. Integrity of evaluation plan implementation
		3.2.2. Appropriateness of remediation and reevaluation implementation
4. Learning outcome		4.1 Degree of satisfaction
	4.1.2. Sufficiency of investigation factors to learner's satisfaction	

Evaluation areas	Evaluation items	Evaluation indicators
		with education
	4.2 Competence achievement	4.2.1. Sufficiency of competence achievement in learner's knowledge, skills and attitude
		4.2.2. Appropriateness of unintended learning outcome
	4.3 Transfer of learning	4.3.1. Sufficiency of applying basic medical education to clinical medical education
		4.3.2. Sufficiency of applying previous learning contents to clinical practice
		4.3.3. Sufficiency of application to medical sites after graduation
		4.3.4. Change in training course
5. Continuous quality improvement (CQI) of curriculum	5.1 Continuous curriculum quality assurance and improvement	5.1.1. Appropriateness of evaluation monitoring system for curriculum improvement
		5.1.2. Appropriateness of curriculum evaluation method
		5.1.3. Sustainability of success factors of curriculum
		5.1.4. Feedback of curriculum evaluation results

Discussion

Experts from nearly 50% of the medical schools in Korea participated in the preparation of the evaluation criteria and the verification of their validity. Therefore, the significance of the present study is that the validity of the developed evaluation tool was verified by the experts throughout Korea.

The validity of the evaluation criteria selected by the literature survey and the expert interview was verified

by the Delphi methodology, which showed that most of the evaluation criteria are valid. In comparison with the first Delphi survey, the consensus, convergence and CV values in the second Delphi survey were found more positive, indicating that the members of the Delphi panel were gradually approaching a consensus. This showed that the characteristics of the Delphi survey were significantly realized in the present study, including the repeated procedures, controlled feedback and the anonymity of the Delphi panel members.

Since most of the drafted evaluation criteria were proved to be valid, it was found that the CBC of medical schools needs to be evaluated by using the evaluation criteria in the extensive range of ‘context, input, process, learning outcome and CQI.’ This has also verified the validity of the mixed model prepared by combining the CIPP evaluation model and the Kirkpatrick evaluation model. In particular, the validity of the evaluation areas, ‘2. Input’ and ‘1. Context,’ was highest. Since the number of valid evaluation indicators included in these two areas is highest, evaluation needs to be carried out by focusing on the areas.

The review of the evaluation criteria that were dropped out from the validation process indicates that most of them involved difficulties in the practical application. ‘4.4 Performance accomplishments by graduates after graduation,’ ‘4.4.2 Change in patients and treatment,’ and ‘4.4.3 Change in society and science,’ the evaluation criteria that were dropped out, correspond to Level 4 (‘results’) of the Kirkpatrick evaluation model, and the limitations in the actual evaluation have been pointed out in many previous reports (62) (28). Although their validity was not proved, the evaluation criteria were prepared in the evaluation criteria drafting stage, and thus they may be evaluated in the future.

In addition, ‘3.3.3 Appropriateness of competence acquisition opportunity in clinical practice’ includes the long-dated controversy in the field of medical education in Korea, such as ‘How can it be promoted to provide more opportunities for medical school students to acquire necessary competences by meeting real patients in clinical settings?’ (63).

Study limitations

This study has the following research limitations. First, this study derived evaluation criteria based on previous research, interviews with 5 experts, and Delphi survey with 20 participants. Therefore, it is difficult to say that the presented evaluation criteria reflected all the reality of competency-based curriculum in domestic medical schools.

Second, this study drew the evaluation criteria by collecting the opinions of experts. However, there are various stakeholders in medical education. For this reason, the results of this study are limited in that they do not reflect the opinions of various stakeholders, such as students, other than experts.

Conclusion

The significance of the present study is that an evaluation model and its evaluation criteria suitable for the curriculums of the medical schools in Korea were developed with the consensus of the experts. The preparation of valid evaluation criteria will open up the possibility of improving the evaluation in medical schools, contributing to the improvement of the educational quality and the continued quality assurance of medical education, which are the ultimate goals of the curriculum evaluation.

This study was extracted and organized from the So Young Lee's doctoral thesis in 2021.

Acknowledgements

We would like to thank all the experts who participated in the interviews and Delphi surveys for this study.

Funding

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF- 2020S1A5B5A17089960)

Abbreviations

CBC: Competency-Based Curriculum

CIPP: Context, Input, Process and Product

CVR: Content Validity Ratio

CV: Coefficient of Variation

CQI: Continuous quality improvement

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

The study protocol was performed in accordance with the Declaration of Helsinki, reviewed and approved by the institutional research board of the Seoul National University College of Medicine (IRB No. C-2005-111-1123).

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Authors' contributions

SYL and SHL conceptualized the study design, and drafted the survey questionnaire. SYL, SHL, JSS reviewed the draft and questionnaire. SYL acquired the data and wrote the initial manuscript draft. SHL, JSS, JAK, HBY critically reviewed initial manuscript draft. According to the reviews, SYL revised the draft. All authors read and approved the final submitted manuscript.

Corresponding author

Correspondence to Seung-Hee Lee

Authors' information

Department of medical education Chungnam National University College of Medicine, Daejeon, Korea

So Young Lee

Department of Medical Education Seoul National University College of Medicine, Seoul, Korea

Seung-Hee Lee, Jwa-Seop Shin

Department of Human Systems Medicine Seoul National University College of Medicine, Seoul, Korea

Seung-Hee Lee, Jwa-Seop Shin, Hyun Bae Yoon

Department of Dental Education in Yonsei University College of Dentistry, Seoul, Korea

Jooah Kim

Office of Medical Education Seoul National University College of Medicine, Seoul, Korea

Hyun Bae Yoon

ORCID

So Young Lee <https://orcid.org/0000-0002-7268-2261>

Seung-Hee Lee <https://orcid.org/0000-0001-8672-5253>

Jwa-Seop Shin <https://orcid.org/0000-0002-6251-3616>

Jooah Kim <https://orcid.org/0000-0003-1307-4239>

Hyun Bae Yoon <https://orcid.org/0000-0003-4367-5350>

References

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *Jama*. 2002;287(2):226-35.
2. Jung H-Y, Kim J-W, Lee S, Yoo SH, Jeon J-H, Kim T-w, et al. A study of core humanistic competency for developing humanism education for medical students. *Journal of Korean medical science*. 2016;31(6):829.
3. Kim SC, Ro YS, Do Shin S, Wi DH, Jeong J, Park JO, et al. Assessment of competence in emergency medicine among healthcare professionals in Cameroon. *Journal of Korean medical science*. 2017;32(12):1931.
4. Magnusson K, Osborne J. The rise of competency-based education: A deconstructionist analysis. *The Journal of Educational Thought (JET)/Revue de la Pensée Educative*. 1990:5-13.
5. Harden RM. AMEE Guide No. 14: Outcome-based education: Part 1-An introduction to outcome-based education. *Medical teacher*. 1999;21(1):7-14.
6. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Medical teacher*. 2010;32(8):631-7.
7. Lurie SJ. History and practice of competency-based assessment. *Medical education*. 2012;46(1):49-57.
8. Shin J, Lee Y, Park W, Baik N, Hwang Y, Yoo K, et al. Development of educational curriculum in medical school. Seoul: Seoul National University College of Medicine. 2008.
9. Lee YM, So YH, Ahn DS, Park SH. Evaluating clinical teaching in medicine. *Korean Journal of Medical Education*. 2002;14(2):293-300.
10. Park JH, Kim S. Evaluating clinical clerkship in medical school. *Korean Journal of Medical Education*. 2004;16(2):157-67.

11. An JH, Han JJ, Eo E, Kwon I, Lee SN. Analytical case study of evaluation of curriculum at a medical school. Korean journal of medical education. 2010;22(1):57-64.
12. Im EJ, Chang B-H, Im E-J, Chang B-H. Analysis on teaching evaluation of clinical clerkship at Korean medical schools. Korean journal of medical education. 2011;23(2):119-26.
13. Fitzpatrick JL. Program evaluation alternative approaches and practical guidelines 2011.
14. Worthen BR, Sanders JR, Fitzpatrick JL. Program evaluation: Alternative approaches and practical guidelines: Longman; 1997.
15. Russ-Eft D, Preskill H. Evaluation in organizations: A systematic approach to enhancing learning, performance, and change: Basic Books; 2009.
16. Cook DA. Twelve tips for evaluating educational programs. Medical teacher. 2010;32(4):296-301.
17. Dorri S, Akbari M, Sedeh MD. Kirkpatrick evaluation model for in-service training on cardiopulmonary resuscitation. Iranian journal of nursing and midwifery research. 2016;21(5):493.
18. Al-Khathami AD. Evaluation of Saudi family medicine training program: The application of CIPP evaluation format. Medical teacher. 2012;34(sup1):S81-S9.
19. Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE guide no. 67. Medical teacher. 2012;34(5):e288-e99.
20. Yarmohammadian MH, Mohebbi N. Review evaluation indicators of health information technology course of master's degree in medical sciences universities' based on CIPP Model. Journal of education and health promotion. 2015;4.
21. Neyazi N, Arab M, Farzianpour F, Mahmoudi M. Identifying weaknesses in undergraduate programs within the context input process product model framework in view of faculty and library staff in 2014. Korean journal of

medical education. 2016;28(2):185.

22. Mirzazadeh A, Gandomkar R, Hejri SM, Hassanzadeh G, Koochak HE, Golestani A, et al. Undergraduate medical education programme renewal: a longitudinal context, input, process and product evaluation study.

Perspectives on medical education. 2016;5(1):15-23.

23. Rooholamini A, Amini M, Bazrafkan L, Dehghani MR, Esmailzadeh Z, Nabeiei P, et al. Program evaluation of an integrated basic science medical curriculum in Shiraz Medical School, using CIPP evaluation model. Journal of Advances in Medical Education & Professionalism. 2017;5(3):148.

24. Ashghali-Farahani M, Ghaffari F, Hoseini-Esfidarjani S-S, Hadian Z, Qomi R, Dargahi H. Neonatal intensive care nursing curriculum challenges based on Context, Input, Process, and Product evaluation model: a qualitative study. Iranian journal of nursing and midwifery research. 2018;23(2):111.

25. Barr H, Freeth D, Hammick M, Koppel I, Reeves S. Evaluations of interprofessional education: a United Kingdom review for health and social care. 2000.

26. Belfield C, Thomas H, Bullock A, Eynon R, Wall D. Measuring effectiveness for best evidence medical education: a discussion. Medical teacher. 2001;23(2):164-70.

27. Tian J, Atkinson NL, Portnoy B, Gold RS. A systematic review of evaluation in formal continuing medical education. Journal of Continuing Education in the Health Professions. 2007;27(1):16-27.

28. Kirkpatrick D, Kirkpatrick J. Evaluating training programs: The four levels: Berrett-Koehler Publishers; 2006.

29. Yoon HB, Shin J-S, Lee S-H, Kim D-H, Hwang J, Kim EJ, et al. The effect of formative program evaluation on continuous program improvement: a case study of a clinical training program in Lao PDR. Journal of Korean medical science. 2015;30(12):1743.

30. Cho D, Cosimini M, Espinoza J. Podcasting in medical education: a review of the literature. Korean journal of

medical education. 2017;29(4):229.

31. Kim D-H, Lee J-H, Kim SA. The pharmacology course for preclinical students using team-based learning. Korean journal of medical education. 2020;32(1):35.

32. Hodge S. The Origins of Competency-Based Training. Australian journal of adult learning. 2007;47(2):179-209.

33. Ten Cate O, Billett S. Competency-based medical education: origins, perspectives and potentialities. Medical Education. 2014;48(3):325-32.

34. Schunk DH. Learning theories an educational perspective sixth edition: Pearson; 2012.

35. Kirkpatrick DL. Implementing the four levels: A practical guide for effective evaluation of training programs: Easyread super large 24pt edition: ReadHowYouWant. com; 2009.

36. SY Kim, YM Lee, YS Choi, SM Hong. Evaluation Research on the Outcomes of Consortium Training Programs for Small and Middle-sized Corporations by Using CIPP model and Kirkpatrick's Model .The Journal of Vocational Education Research. 2011;30(1):51-74.

37. DH Kim. Development of strategies for implementation and evaluation of an international faculty development program based on outcome-oriented and process-oriented research .PhD thesis, Seoul National University, College of Medicine (2018)

38. YJ Koo·WS Chang. A Cultural Historical Activity Theoretical Approach on Highly Educated Women 'Becoming Working Mothers .The Women's Studies. 2018;98(3):5-34.

39. DS Lee, YC Kim. An Inquiry on the Inclusive Analytic Procedures for Qualitative Data Analysis: Based on the Pragmatic Eclecticism .The Journal of Educational Research. 2014;12(1):159-84.

40. Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning: A guide to nominal

group and Delphi processes: Scott, Foresman; 1975.

41. Dalkey NC. Information Systems in Management Science—A Case Study of a Decision Analysis: Hamlet's Soliloquy. *Interfaces*. 1981;11(5):45-9.

42. Saaty TL. Fundamentals of decision making and priority theory with the analytic hierarchy process: RWS publications; 2000.

43. Lawshe CH. A quantitative approach to content validity. *Personnel psychology*. 1975;28(4):563-75.

44. JS Lee. *Delphi Method*. Seoul: Kyowookbook. 2001;7.

45. YJ Jeon. Study on The Job Performance Competencies Criteria of Technology Teacher Related Curriculum Education. *KJTE*. 2006;6(1):121-55.

46. YJ Kang. Understanding and application of Delphi techniques. Korea Employment Agency for the Disabled: project report. 2008:1-17.

47. YH Khang, SJ Yoon, GW Kang, CY Kim, KY Yoo, YS Shin. An Application of Delphi Method to the Assessment of Current Status of Cancer Research. *J Prev Med Public Health*. 1998;4(31):844-56.

48. SY Noh. The Delphi Technique: Forecasting the Future with Expert Insight. Korea Research Institute For Human Settlements: PLANNING AND POLICY. 2006:53-62.

49. Lee YM, Rhee KJ, Lee YJ, Ahn DS. Evaluation of Integrated Lectures at Korea University, College of Medicine. *Korean J Med Educ*. 2000;12(2):343-51.

50. JH Park, S Kim. Evaluating Clinical Clerkship in Medical School. *Korean J Med Educ*. 2004;16(2): 157-167

51. JH An, JJ Han, Nj Kim, EK Eo, I Kwon and SN Lee. Analytical Case Study of Evaluation of Curriculum at a Medical School. *Korean J Med Educ*. 2010;22(1):57-64.

52. JY Hwang. Research and Development of Curriculum Evaluation Criteria for Premedical Curriculum .Master Degree thesis, Seoul National University, College of Medicine (2015)
53. National Teacher Training Center for Health Personnel, Seoul National University College of Medicine. Educational Program Evaluation Guide for Transforming Medical School Education .National Teacher Training Center for Health Personnel, Seoul National University College of Medicine. 2019
54. Cha BE · Shon MH. A Study on the Development of Evaluation Indicators for the Education Process in Medical School .JOEC. 2019;25(4):239-61.
55. Lee S, Shin JS, Kim D-H, Kim EJ, Hwang J, Yoon HB, et al. Evaluation of first year outcomes of Lee Jong-Wook Seoul Project using Kirkpatrick's four-level evaluation approach. Korean journal of medical education. 2014;26(1):41-51.
56. Abdulghani HM, Shaik SA, Khamis N, Al-Drees AA, Irshad M, Khalil MS, et al. Research methodology workshops evaluation using the Kirkpatrick's model: translating theory into practice. Medical Teacher. 2014;36(sup1):S24-S9.
57. Simpson JS, Scheer A. A review of the effectiveness of breast surgical oncology fellowship programs utilizing Kirkpatrick's evaluation model. Journal of Cancer Education. 2016;31(3):466-71.
58. Lee SY, Lee S-H, Shin J-S. Evaluation of medical humanities course in college of medicine using the Context, Input, Process, and Product evaluation model. Journal of Korean medical science. 2019;34(22).
59. Heydari MR, Taghva F, Amini M, Delavari S. Using Kirkpatrick's model to measure the effect of a new teaching and learning methods workshop for health care staff. BMC research notes. 2019;12(1):1-5.
60. YS Im, WS Choi, SC Song. A Study on the Evaluation of the University Specialized-degree Program by Applying Kirkpatrick's Model .KJMAS. 2013;69(1):131-63.

61. Min, KS · Choe DH · Moon KH. A Study on the Indicators for Evaluating Police Training and Learning .PSJ. 2015;10(2):9-33.
62. Jung JS. Education program evaluation .Seoul: Kyowookbook. 2014
63. Kim HJ, Huh JS. The Right of the Clinical Training for the Medical Students and Privacy of the Patients .Korean Journal of Medicine and Law. 2013;21(2):107-23.

Appendix

[Appendix 1] Evaluation indicators derived from the first and second Delphi surveys.

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
1.1.1. Clarity of educational competences and educational goals	4.6	1	0.8	0.5	0.11	1.1.1. Clarity of educational competences and educational goals	4.7	1	0.8	0.5	0.1
1.1.2. Appropriateness of analyzing stakeholders' demands in educational competences and	4.25	0.7	0.75	0.5	0.2	1.1.2. Appropriateness of analyzing stakeholders' demands in educational competences and	4.4	0.9	0.75	0.5	0.14

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
education goals						education goals					
1.1.3. Sufficiency of reflecting essential medical education competences to educational competences an education goals	4.53	0.68	0.8	0.5	0.17	1.1.3. Sufficiency of reflecting essential medical education competences to educational competences an education goals	4.7	1	0.85	0.38	0.1
1.1.4. Clarity of recognition of educational competences and educational goals	4.05	0.6	0.75	0.5	0.2	1.1.4. Clarity of recognition of educational competences and educational goals	4.1	0.7	0.94	0.13	0.16
1.3.1. Integrity of	4.45	0.7	0.8	0.5	0.17	1.2.1. Integrity of	4.7	0.9	1	0	0.12

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
learner analysis						learner analysis					
1.3.2. Integrity of teacher analysis	4.4	0.8	0.78	0.5	0.15	1.2.2. Integrity of teacher analysis	4.7	1	0.85	0.38	0.1
1.4.1 Sufficiency of the will and leadership of executives	3.8	0.4	-	-	-	-	-	-	-	-	-
1.4.2 Clarity of finance and regulations for curriculum implementation	4.55	0.9	0.8	0.5	0.13	1.3.1 Integrity of finance and regulations for curriculum implementation	4.74	1	1	0	0.1
1.4.3. Sufficiency of member consensus	4.15	0.6	0.75	0.5	0.21	1.3.2. Sufficiency of member consensus	4.5	0.9	0.8	0.5	0.17

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
on curriculum introduction						on curriculum introduction					
1.2.1. Sufficiency of reflecting curriculum design principles	4.2	0.6	0.75	0.5	0.21	2.1.1. Sufficiency of reflecting curriculum design principles	4.5	1	0.78	0.5	0.11
2.2.7. Scrupulosity of syllabus	4.2	0.7	0.75	0.5	0.17	2.1.2 Scrupulosity of syllabus	4.25	1	0.94	0.13	0.1
2.2.8. Appropriateness of amount of learning goals	4.05	0.6	0.75	0.5	0.2	2.1.3 Appropriateness of amount of learning goals	3.9	0.8	1	0	0.2
2.1.1. Appropriateness of	4.6	1	0.8	0.5	0.11	2.2.1. Appropriateness of	4.65	0.9	0.85	0.38	0.13

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
educational contents						educational contents					
2.2.4. Appropriateness of basic medicine-clinical medicine education connection	4.35	0.7	0.8	0.5	0.2	2.2.2. Appropriateness of basic medicine-clinical medicine education connection	4.6	0.9	0.8	0.5	0.13
2.2.6. Appropriateness of balance coordination in curriculum and subjects	3.9	0.5	0.88	0.25	0.23	2.2.3 Appropriateness of balance coordination in curriculum and subjects	4.15	0.9	0.75	0.5	0.21
2.1.2. Appropriateness of education methods	4.65	1	0.8	0.5	0.11	2.3.1. Appropriateness of education methods	4.6	0.9	0.8	0.5	0.13

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
2.2.1. Appropriateness of educational programming for clinical setting preparation	4.25	0.7	0.75	0.5	0.2	2.3.2. Appropriateness of educational programming for clinical setting preparation	4.05	0.7	0.75	0.5	0.23
2.2.2. Appropriateness of clinical educational programming in consideration of learner's level and education after graduation	4.05	0.5	0.69	0.63	0.22	2.3.3. Appropriateness of clinical educational programming in consideration of learner's level and education after graduation	3.85	0.7	1	0	0.21

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
2.2.3. Sufficiency of providing education opportunities through meetings with patients	4.25	0.7	0.75	0.5	0.17	2.3.4. Sufficiency of providing education opportunities through meetings with patients	4.2	0.9	0.94	0.13	0.12
	4.2	0.6	0.78	0.5	0.24	2.3.5 Practice reinforcement	4.25	0.8	0.75	0.5	0.23
2.3.1. Clarity of establishing basic principles of evaluation	4.45	0.8	0.8	0.5	0.15	2.4.1. Clarity of establishing basic principles of evaluation	4.85	1	1	0	0.08
2.3.2. Appropriateness of evaluation methods	4.7	1	0.8	0.5	0.1	2.4.2. Appropriateness of evaluation methods	4.65	0.9	0.85	0.38	0.13

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
2.3.3. Evaluation focused on performance	4.35	0.7	0.78	0.5	0.17	2.4.3. Evaluation focused on performance	4.45	0.9	0.8	0.5	0.21
2.3.4. Clarity of competence achievement evaluation criteria	4.4	0.9	0.75	0.5	0.14	2.4.4. Clarity of competence achievement evaluation criteria	4.45	1	0.75	0.5	0.11
2.3.5. Appropriateness of introducing progress test	3.95	0.5	0.88	0.25	0.21	2.4.5. Appropriateness of introducing progress test	4	0.9	1	0	0.2
2.3.6. Appropriateness of introducing absolute	3.58	0.16	-	-	-	-	-	-	-	-	-

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
evaluation'											
2.3.7. Appropriateness of providing remediation and reevaluation	4.4	0.9	0.75	0.5	0.14	2.4.6. Appropriateness of providing remediation and reevaluation	4.4	0.9	0.75	0.5	0.14
2.3.8. Appropriateness of portfolio evaluation	3.65	0.5	0.94	0.13	0.27	2.4.7. Appropriateness of portfolio evaluation	3.95	0.8	1	0	0.1
2.3.9. Appropriateness of formative evaluation and summative evaluation	4.3	0.6	0.78	0.5	0.19	2.4.8. Appropriateness of formative evaluation and summative evaluation	4.7	1	0.85	0.38	0.1

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
2.4.1. Clarity of granting authorities and roles to curriculum supervising professor	4.3	0.7	0.75	0.5	0.17	2.5.1. Clarity of granting authorities and roles to curriculum supervising professor	4.6	1	0.8	0.5	0.11
2.4.2. Integrity of organization in charge of medical education	4.25	0.8	0.75	0.5	0.19	2.5.2. Integrity of organization in charge of medical education	4.3	1	0.75	0.5	0.11
2.4.3. Integrity of organization in charge of evaluation'	4.25	0.7	0.75	0.5	0.17	2.5.3. Integrity of organization in charge of evaluation'	4.45	1	0.75	0.5	0.11
2.4.4.	4.4	0.7	0.8	0.5	0.17	2.5.4.	4.6	0.9	0.8	0.5	0.13

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
Appropriateness of support for teaching capability development						Appropriateness of support for teaching capability development					
2.6.1. Appropriateness of learning support for learners	4.35	0.9	0.75	0.5	0.13	2.6.1. Appropriateness of learning support for learners	4.37	1	0.75	0.5	0.11
2.5.1. Sufficiency of facilities and equipment	4.4	0.8	0.78	0.5	0.15	2.7.1. Sufficiency of facilities and equipment	4.55	0.9	0.8	0.5	0.13
2.5.2. Appropriateness of cooperation between	4.05	0.5	0.69	0.63	0.22	2.7.2. Appropriateness of cooperation between	4.25	0.9	0.75	0.5	0.21

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
medical schools and hospitals						medical schools and hospitals					
3.1.2. Integrity of education plan implementation	4.37	0.79	0.75	0.5	0.16	3.1.1 Integrity of education plan implementation	4.26	1	0.88	0.25	0.11
3.1.3. Smoothness of conducting education	4.25	0.7	0.75	0.5	0.17	3.1.2 Smoothness of conducting education	4.21	1	1	0	0.1
3.1.4. Appropriateness of feedback	4.5	0.8	0.8	0.5	0.18	3.1.3. Appropriateness of feedback	4.68	1	0.8	0.5	0.1
3.2.1. Integrity of evaluation plan	4.35	0.7	0.8	0.5	0.2	3.2.1. Integrity of evaluation plan	4.74	1	1	0	0.1

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
implementation						implementation					
3.2.2. Appropriateness of remediation implementation	4.4	0.7	0.8	0.5	0.17	3.2.2. Appropriateness of remediation implementation	4.53	0.89	0.8	0.5	0.14
3.3.1. Appropriateness of education evaluation	4.4	0.7	0.8	0.5	0.17	3.3.1. Appropriateness of education evaluation method	4.74	0.89	1	0	0.12
3.3.2. Collection of education evaluation results	4.4	0.8	0.8	0.5	0.19	3.3.2. Collection of education evaluation results	4.63	0.89	0.8	0.5	0.13
3.3.3. Sustainability of the activities by	3.95	0.6	0.75	0.5	0.14	3.3.3 Sustainability of the activities by	4.32	0.89	0.75	0.5	0.13

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
the competence review committee						the competence review committee, the curriculum development committee, and the education evaluation committee					
3.3.4. Appropriateness of curriculum supervising professor'	4.1	0.6	-	-	-	-	-	-	-	-	-
4.1.1. Sufficiency of learner's satisfaction	4.35	0.9	0.75	0.5	0.17	4.1.1. Sufficiency of learner's satisfaction	4.15	0.9	0.75	0.5	0.21

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
with education						with education					
4.1.2. Sufficiency of investigation factors to learner's satisfaction with education	4.1	0.7	0.75	0.5	0.19	4.1.2. Sufficiency of investigation factors to learner's satisfaction with education	3.95	0.8	1	0	0.24
4.2.1. Sufficiency of competence achievement in learner's knowledge, skills and attitude	4.55	1	0.8	0.5	0.11	4.2.1. Sufficiency of competence achievement in learner's knowledge, skills and attitude	4.75	1	1	0	0.09
4.2.2. Appropriateness of unintended learning	3.95	0.7	1	0	0.21	4.2.2. Appropriateness of unintended learning	4.05	0.9	1	0	0.1

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
outcome						outcome					
4.3.1. Sufficiency of applying basic medical education to clinical medical education	4.2	0.8	0.75	0.5	0.18	4.3.1. Sufficiency of applying basic medical education to clinical medical education	4.15	0.8	0.94	0.13	0.14
4.3.2. Sufficiency of applying previous learning contents to clinical practice	4.1	0.6	0.75	0.5	0.21	4.3.2. Sufficiency of applying previous learning contents to clinical practice	4.2	0.8	0.75	0.5	0.15
4.3.3. Sufficiency of application to medical sites after graduation	4.25	0.8	0.75	0.5	0.19	4.3.3. Sufficiency of application to medical sites after graduation	4.15	0.8	0.75	0.5	0.22

First Delphi surveys						Second Delphi surveys					
Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV	Evaluation item	Mean	CVR	Degree of consensus	Degree of convergence	CV
4.4.1. Changes during resident course	3.85	0.5	0.94	0.13	0.19	4.4.1. Improvement of competences during resident course	3.8	0.6	1	0	0.22
5.1.1. Sustainability of success factors of curriculum	4.05	0.7	0.94	0.13	0.19	5.1.1. Sustainability of success factors of curriculum	4	0.8	1	0	0.21
5.1.2. Appropriateness of feedback system for curriculum improvement	4.5	1	0.78	0.5	0.11	5.1.2. Appropriateness of feedback system for curriculum improvement	4.85	1	1	0	0.08

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20211102EvalCriteriaCBCMedicalSchoolappendix.pdf](#)