Peer review analysis of lectures using video recordings in an integrated curriculum

Dong-Mi Yoo¹, HyunBae Yoon² and Seunghee Lee²

¹National Teacher Training Center for Health Personnel and ²Department of Medical Education, Seoul National University College of Medicine, Seoul, Korea

통합교육과정에서 녹화된 강의 동영상을 활용한 동료평가 분석 연구

서울대학교 의과대학 ¹의학교육연수원, ²의학교육학교실

유동미¹, 윤현배², 이승희²

Purpose: The purpose of this study is to evaluate the effect of the peer review in an integrated curriculum and to guide further improvements of curriculum.

Methods: In 2012, Seoul National University College of Medicine implemented a peer review system for 11 courses in an integrated curriculum. For each lecture, two reviewers conducted the rating using a 10-item questionnaire on a 4-point scale. We analyzed the correlation between total scores and each item and the inter-rater reliability between the two reviewers by Pearson correlation. Further, the link between peer review scores and the student lecture evaluation was analyzed.

Results: The mean total score for the checklist rating was 31.3 (out of 40.0), and the mean score for each item ranged from 2.65 to 3.35 (out of 4.00). The correlation coefficient between the total score and each item was high, ranging from 0.656 to 0.849, except for three items. The mean of difference scores between reviewers was 5.03, and the correlation coefficient was significantly high, which ranged from 0.968 to 0.999. The peer reviews scores and student lecture evaluations generally correlated, but there were some outlying exceptions; the correlation coefficient was 0.105 and 0.093.

Conclusion: Peer review is a useful method for improving the quality of lectures in an integrated curriculum by monitoring the objectives, contents, and methods of the lectures and providing feedback to the professors.

Key Words: Peer review, Integrated curriculum, Lecture, Video recording

Received: October 29, 2014 • Revised: December 24, 2014 • Accepted: January 20, 2015

Corresponding Author: Seunghee Lee (<http://orcid.org/0000-0001-8672-5253>)

Department of Medical Education, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 110-799, Korea

Tel: +82.2.740.8406 Fax: +82.2.745.6373 email: lscho@snu.ac.kr

Korean J Med Educ 2015 Mar; 27(1): 11-18.
<http://dx.doi.org/10.3946/kjme.2015.27.1.11>
eISSN: 2005-7288

© The Korean Society of Medical Education. All rights reserved.
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서론

통합교육과정의 개념은 1950~60년대 미국에서 학문단위로 세분화된 교과목 모두를 나열식으로 배열하여 학습시키기 보다는 서로 관련 있는 것끼리 묶어서 가르치면 배움을 보다 효과적으로 촉진시킬 수 있다는 그간의 연구에 기반을 두어 발생하였다[1]. 이러한 개념을 의학교육에 실제적으로 도입한 것은 1950년 미국 Western Reserve 의과대학이 'inter-departmental teaching'이라는 형식으로 교육과정을 개편하면서부터이다[2]. 우리나라도 1971년에 서울대학교 의과대학에서 'block lecture'라는 이름으로 처음 통합교육을 실시하였고, 이후 2000년도에 들어서는 전국 41개 대학 중 35개 대학이 통합교육과정을 시행하고 있으며, 의과대학 인증평가에서도 통합교육을 권장하고 있어 형식적으로는 거의 모든 의과대학이 통합교육을 시행하고 있다고 할 수 있다[3,4].

의학교육에서 통합교육과정은 흩어진 단편적인 지식들을 불필요한 반복 없이 일정 시기에 모아서 논리적으로 배열하여 여러 전문교수가 함께 가르침으로써 학생들의 학습에 도움을 주는 장점이 있다. 통합교육이 제대로 이루어지기 위해서는 타 학과나 동일 교실 내의 다른 교수와의 능동적인 협조가 필수적인데, 이 부분이 제대로 실행되지 않으면 강의의 체계성 결여, 통합강의의 일관된 목표 부재, 과목중심강의와의 차별이 없어지는 등의 제한점을 가질 수밖에 없기 때문에 통합교육을 시행한 후에는 교육의 수월성을 재검토할 필요가 있다[5].

교육의 질적 수준 제고를 위한 노력의 일환으로 수업평가의 중요성에 대한 인식이 확산되면서 많은 대학에서 강의평가가 시행되고 있으나, 대부분 학생들의 강의평가 혹은 강좌평가에 의존하고 있으며, 또 다른 강의평가의 유용한 방법 중 하나인 동료평가(peer review)는, 수업의 공개와 평가를 부담스러워하는 우리나라 대학 교육 풍토에서 거의 시행되지 못하고 있는 실정이다[6].

이러한 동료평가는 교육활동의 실제 주체인 교수가 현장에서 직접 접하는 동료교수의 교수수행능력을 관찰하고 평가한다는 점에서 그 결과를 교수 스스로 신뢰할 가능성이 크며, 수업을 공개하는 피평가자뿐만 아니라 이를 검토하는 평가자의

발전도 기대할 수 있는 효과적인 방법이라 할 수 있다[6,7].

또한, 의과대학에서의 동료평가 방법은 수업의 질을 향상시키고, 교수들이 건설적으로 자기평가를 하도록 유도하며, 교수들의 의학교육과정에 대한 이해도를 향상시킴으로써 교육과정 개선에 적극 참여하게 하며, 우수한 수업과 그렇지 못한 수업에 대한 정의를 분명히 하는 데 도움을 주어 교육과정 개선 시 유용한 정보를 제공해 주는 긍정적인 역할을 한다[8,9].

국내 의과대학에서는 대부분 학생들에 의한 강의평가는 이뤄지고 있으나 이에 대한 직접적인 연구 결과는 많지 않으며, 주로 우수 강의자의 특성이나 좋은 수업의 요건에 관한 연구가 대부분이다[10,11]. 이러한 학생 설문 중심의 강의평가는 평가자인 학생의 전문적 판단수준 미흡, 응답에 대한 성실성의 부재, 학생 만족도 중심의 설문 평가문항에 따른 실제 교수 역량 측정의 미흡, 강의 외적인 요인의 영향 등 평가의 타당도와 신뢰도에 대한 문제를 깔끔히 해소시키지는 못하고 있는 것이 현실이다[12,13]. 통합교육과정의 평가에 관한 연구도 지속적으로 있었으나, 대부분 학생과 교수들을 대상으로 한 설문조사에 기반을 둔 연구들이었다[14,15].

더불어, 그 동안 의학교육에서 강의의 동영상 녹화는 대체로 학생들이 복습을 하거나 듣지 못한 강의를 보충하기 위한 목적으로 이용되어 왔다[16]. 하지만, 의과대학에서 교육의 많은 부분이 여전히 강의로 이루어지고 있기 때문에 동영상을 활용한 강의평가는 의학교육의 질을 향상시키는 데 기여할 가능성이 높다. 특히, 녹화된 강의 동영상을 활용한다면 물리적, 시간적 제약으로부터 비교적 자유롭게 동료평가를 수행할 수 있다는 장점이 있다. 실제로 최근 국내 한 의과대학에서 비록 소수의 교수에 대하여 제한적으로 시행하기는 하였으나, 녹화된 강의 동영상을 활용하여 수업을 분석한 결과를 보고한 바가 있다[17].

이러한 배경에서 서울대학교 의과대학에서는 2010년 11월부터 2011년 2월까지 기존의 전체 교육과정에 대한 평가를 진행하였고, 이어서 2011년 2월부터 4월까지 통합교육과정에 대한 시범적인 동료평가가 진행되었다. 시범적인 동료평가는 강의록을 이용하여 간접적으로 이루어졌는데, 강의록의 내용이 너무 방대하고 수준이 너무 높은 것으로 나타났다. 이에 2012년에는 통합교육과정 전 과목의 모든 강의를 동영상으로

촬영하여 이를 바탕으로 동료평가를 시행하기로 하였다. 본 연구에서는 일개 의과대학의 통합교육과정에서 녹화된 강의 동영상상을 활용한 동료평가를 분석하여 동료평가의 효과와 의의를 평가하고 향후 보완해야 할 점들을 제시하고자 한다.

대상 및 방법

1. 통합교육과정에서의 동료평가 방법 개요

2012년도에 서울대학교 의학과 2학년 학생들을 대상으로 진행된 통합교육과정의 11개 과목(감염학, 임상면역학, 의학유전학, 중양학, 내분비학, 혈액학, 신장요로학, 순환기학, 호

흡기학, 소화기학, 신경계학)의 모든 강의에 대하여 동료평가를 시행하였다. 강의하는 교수들에게 동영상 촬영 및 동료평가에 대하여 사전에 안내를 시행한 후에 모든 강의를 촬영하여 동영상상으로 제작하였다. 전체 교수들을 대상으로 평가자(reviewer) 신청을 받아서 평가를 진행할 강의를 배분하였다. 총 435명의 교수가 진행한 561시간의 강의에 대하여 421명의 교수가 동료평가를 시행하였다. 각 수업마다 2명의 평가자가 녹화된 동영상상을 보면서 4점 척도의 10개 문항으로 구성된 체크리스트를 활용하여 평가를 하였고 서술형 의견도 기술하게 하였다.

체크리스트의 문항은 강의의 목적, 수준의 적절성, 분량의 적절성, 학습 목표의 유무, 목표와 내용의 일치, 사례의 제시, 명확한 언어 전달, 학생들과의 상호작용, 핵심 요약, 전반적인 만족도 등으로 구성되었다(Table 1). 동료평가가 완료된 후에는 강의한 교수에게 강의 동영상상과 체크리스트, 그리고 평가결과와 피드백 등을 제공하였다(Fig. 1).

한편, 학생들을 대상으로 각 과목 종료 후 강의평가를 시행하였으며, 그 중에서 강의 분량의 적절성과 교수법이 흥미롭고 전달력이 좋았는지를 묻는 2개 문항은 교수에 의한 동료평가 항목과 동일하게 구성되었다.

2. 연구 방법 및 자료 분석

첫째, 각 과목별 평가 결과와 체크리스트의 각 항목별 평가결과를 비교 분석하였으며, 체크리스트의 각 항목과 총점의 상관관계, 그리고 각 항목 간의 상관관계를 분석하였다. 둘째,

Fig. 1. The Process of Peer Review in an Integrated Curriculum

Table 1. Mean and Standard Deviation of Each Item in the Peer Review

Item	Mean	SD
The class is aligned with the goals of the integrated curriculum.	3.35	0.60
The class is appropriate for the student level.	3.24	0.69
The class fits into 50 minutes.	3.11	0.94
Learning outcomes are announced.	3.16	0.94
The content of the class aligns with the learning outcome.	3.26	0.75
Appropriate cases and examples are used to effect students' understanding.	3.22	0.72
The lecturer has good linguistic communication skills (pronunciation, speed, volume).	3.32	0.70
There is interaction between the lecturer and students including questions and answers.	2.65	0.84
Before the end of class, there is a summary of the key points of the class.	2.78	0.92
As a whole, the class is satisfactory.	3.19	0.67

SD: Standard deviation.

각 강의마다 두 명의 평가자가 강의를 평가한 결과 간의 Pearson의 상관계수와 평가자 간의 차이점수를 산출함으로써 평가자 간 신뢰도(inter-rater reliability)를 분석하였다. 셋째, 교수에 의한 동료평가 항목과 동일한 학생들의 강의평가 항목 간의 상관관계를 분석하였다. 통계분석에는 SPSS version 20.0 (SPSS Inc., Chicago, USA)을 이용하였다.

결과

1. 각 항목별 평가 결과

1) 과목별 평균

체크리스트를 활용한 평가에서는 40점 만점에 평균 점수는 31.3 ± 3.8 점이었다. 25분위수가 28.5점, 50분위수가 31.5점,

75분위수가 34.0점으로 나타났으며, 전체 교수들의 점수는 대체로 정규분포를 보였다(Fig. 2).

2) 체크리스트 항목별 평균

체크리스트 10개 문항의 항목별 평균 점수는 5점 만점에 3.13 ± 0.78 이었으며, 특히 '통합교육의 취지에 맞는 수업을 한다는 항목(3.35 ± 0.60)과 '언어적 측면에서 전달이 잘 된다'는 항목(3.32 ± 0.70)에서 높게 나타났고, '질문과 답변을 이끌어내는 등 학생들과 상호작용을 한다' 항목(2.65 ± 0.84)과 '수업이 끝나기 전 중요한 점을 요약한다' 항목(2.78 ± 0.92)은 낮은 점수를 나타냈다(Table 1).

3) 체크리스트 항목별 상관관계

체크리스트의 각 항목과 총점 간의 상관계수는 0.543에서 0.849까지 대체로 양호한 상관을 보였으며, 모두 통계적으로 유의하였다($p < 0.01$). 특히 '전반적으로 수업이 만족스럽다' 항목과 '학습 목표와 강의 내용이 일치한다' 항목은 상관계수

Fig. 2. Mean, Maximum, Minimum Scores of Each Subject

Table 2. Pearson Correlation Coefficient between Each Item and Total Scores in the Peer Review

Item	p-value	r
The class is aligned with the goals of the integrated curriculum.	<0.01	0.678
The class is appropriate for the student level.	<0.01	0.656
The class fits into 50 minutes.	<0.01	0.541
Learning outcomes are announced.	<0.01	0.675
The content of the class aligns with the learning outcome.	<0.01	0.716
Appropriate cases and examples are used to effect students' understanding.	<0.01	0.556
The lecturer has good linguistic communication skills (pronunciation, speed, volume).	<0.01	0.543
There is interaction between the lecturer and students including questions and answers.	<0.01	0.636
Before the end of class, there is a summary of the key points of the class.	<0.01	0.630
As a whole, the class is satisfactory.	<0.01	0.849

0.849와 0.779로 총점과 높은 상관을 보이는 것으로 나타났다(Table 2). 한편, 각 항목 간의 상관관계는 전반적으로 높지 않았으나, ‘통합교육의 취지에 맞는 수업을 한다’는 항목과 ‘학생들의 수준에 맞는 내용이다’는 항목의 상관관계는 0.564, ‘학습 목표를 제시한다’는 항목과 ‘학습 목표와 강의 내용이 일치한다’는 항목의 상관관계는 0.779로 높게 나타났다($p < 0.01$).

2. 평가자 간 신뢰도 분석 결과

1) Pearsons의 상관계수 분석

11개 과목의 561시간 각각의 강의를 두 명의 평가자가 평가한 결과 간의 상관을 살펴본 결과, 0.968에서 0.999까지 모든 강의에서 매우 높은 수준의 상관을 보였다. 각 과목별로 상관계수의 평균을 산출한 결과, 모든 과목이 0.990에서 0.995로 매우 높은 수준의 상관을 보였다(Table 3).

Table 3. Pearson Correlation Coefficient between Reviewers

Subject	p-value	r
Infection	<0.01	0.991
Clinical immunology	<0.01	0.994
Medical genetics	<0.01	0.995
Oncology	<0.01	0.993
Endocrinology	<0.01	0.995
Hematology	<0.01	0.994
Nephro-urology	<0.01	0.994
Cardiology	<0.01	0.992
Pulmonology	<0.01	0.990
Gastroenterology	<0.01	0.991
Neurology	<0.01	0.993

2) 평가자 간 차이점수 분석

총점 40점을 기준으로 해서 두 평가자 사이 편차의 평균은 5.03 ± 3.78 점이었다. 체크리스트의 각 항목별 편차의 평균은, ‘통합교육의 취지에 맞는 수업을 한다’는 0.60점, ‘학생들의 수준에 맞는 내용이다’는 0.72점, ‘50분 강의에 맞는 분량이다’가 0.75점, ‘학습 목표를 제시한다’가 0.66점, ‘학습 목표와 강의 내용이 일치한다’가 0.71점, ‘이해를 돕기 위해 적절한 사례나 예시를 사용한다’가 0.69점, ‘언어적 측면에서 전달이 잘 된다’가 0.70점, ‘질문과 답변을 이끌어내는 등 학생들과 상호작용을 한다’가 0.74점, ‘수업이 끝나기 전 중요한 점을 요약한다’가 0.71점, ‘전반적으로 수업이 만족스럽다’가 0.66점이었다. 한편, 과목별 평가자 간 총점의 차이점수는 4.1에서 6.1로 나타났다(Fig. 3).

3. 학생 강의평가 결과

1) 교수 동료평가와 학생 강의평가 간 상관관계

동료평가 체크리스트의 ‘50분 강의에 맞는 분량이다’는 항목과 학생들의 강의평가의 ‘강의 분량은 적절한가?’ 항목 간 Pearson의 상관계수는 0.105였으며, 체크리스트의 ‘언어적 측면에서 전달이 잘 된다’는 항목과 학생들의 강의평가의 ‘흥미롭고 전달력이 좋았는가?’ 항목의 상관계수는 0.093으로 나타났다으며, 모두 통계적으로 유의하였다.

2) 교수의 동료평가와 학생 강의평가 서술식 평가 결과 비교

교수 동료평가와 학생 강의평가 결과가 일치하는 경우 중

Fig. 3. Difference Scores between Reviewers of Each Subject

교수평가와 학생 강의평가에서 모두 최고 점수를 받은 교수는 4명이었는데, 이들은 교수와 학생 모두에게 강의자의 말의 속도가 적당하며 어조가 부드럽고 명확하다, 강의 내용이나 분량이 적절하며 강의자가 학생들이 이해하기 쉽게 설명한다, 학생들과 상호작용이 활발하다, 활기차고 적극적이며 열정적인 태도로 수업을 진행한다는 등의 평가를 받았다.

또한 교수평가와 학생 강의평가에서 모두 최저 점수를 받은 교수는 2명이었는데, 서술의견에서 교수, 학생 모두 강의의 분량이 너무 많고 수업시간이 길다, 단순한 질환 등의 나열로 강의 내용이 구성되었다, 내용이 학생들에게 너무 어려운 수준이다, 강의자가 학생들과 눈을 맞추지 않고 상호작용이 거의 없다, 일방적인 지식 전달 위주이다 등의 평가를 받았다.

한편, 일부 강의자에 대해 서술식 평가에서 교수는 다른 강의와 균형이 필요하며 기본적인 설명이 부족하다고 평가한 반면, 학생들은 해당 강의에 대해 명료한 설명과 전달력이 좋았다는 대체로 긍정적인 평가를 하는 등 교수의 동료평가와 학생의 강의평가에서 상이한 결과를 보인 경우도 있었다.

고찰

의학교육의 통합교육이 갖는 장점은 개별적으로 존재하던 의미 없던 지식에 해석과 판단이 동시에 주어져 이것이 입체적인 합성과 분석을 통해 의미 있는 지식으로 전환된다는 점이다. 이러한 통합교육의 특징이 잘 적용되어 운영되기 위해서는 각 과목 책임교수의 적절한 역할 수행을 통한 헌신과 교수들의 협조, 교수들을 대상으로 한 지속적인 교육, 교육의 수월성을 제고하기 위한 개선 노력이 필수적이다[1,3,4].

서울대학교 의과대학에서는 통합교육과정에서 강의록과 강의내용이 일치하는지, 수업분량이 적절한지, 통합교육의 취지에 맞는 수업인지, 학생들에게 이해가 잘 되는 교수법을 활용하고 있는지 등을 확인한 후 차기 년도 수업에 평가 내용을 반영하도록 함으로써, 강의의 질을 향상시키고, 교육의 수월성을 제고하고자 하는 목적으로 강의가 녹화된 동영상상을 활용한 동료평가를 시행하였다. 그 평가 결과를 살펴보면, 우선 체크리스트를 활용한 동료평가에서 전체 평균 점수는 40점 만점에 31.3점으로 양호한 편이었으며, 각 과목의 평균도

30.0~32.2점으로 큰 차이가 없었다. 하지만, 각각의 교수들의 경우에는 18.5~40.0점으로 점수에 상당한 편차가 존재하였으며, 이는 모든 과목에서 마찬가지였다. 전체 교수들의 점수는 대체로 정규분포를 보였는데, 24.0점 미만의 점수를 받은 교수 비율은 전체의 3.7%였다. 이처럼 매우 낮은 점수를 받은 강의들은 대체로 분량이 너무 많거나 학생들에게 너무 어려운 수준이며, 교수와 학생의 상호작용이 없고 일방적인 지식 전달 위주였다는 평가를 받았다. 따라서 이러한 평가를 받은 교수들의 경우 충분한 피드백과 체계적인 교수법 훈련을 통해 강의 질을 향상시키도록 도움을 줄 필요성이 있다. 실제로, 서울대학교 의과대학에서는 동료평가에서 매우 낮은 점수를 받은 교수들을 대상으로 교수법 훈련을 제공하였으며, 이를 이수하지 않을 경우 다음 학기 강의에 제한을 두도록 하였다.

체크리스트의 각 항목별 평균 점수를 살펴보면, 교수들이 대체로 통합교육의 취지에 맞는 수업을 하고 있고 학생들에게 언어적 측면의 전달은 잘 되고 있는 반면, 수업이 끝나기 전에 중요한 점을 요약하는 것에 여전히 소홀하였으며, 특히 수업 중에 학생들의 질문과 답변을 이끌어 내는 것 등의 상호작용이 가장 미흡하였다. 수업 중에 교수와 학생의 상호작용이 부족하다는 점은 많은 선행 연구에서도 지속적으로 제기되고 있는 만큼, 향후 교수법 훈련에서도 가장 중점적으로 강조될 필요가 있겠다.

체크리스트의 모든 항목은 전체 총점과 중등도 이상의 상관관계가 있었다. 특히 학습 목표와 강의 내용 일치 정도에 관한 항목의 결과가 총점과 상관관계가 높아 학습 목표와 내용을 일치시키는 것이 좋은 강의의 핵심적인 요소라는 사실을 다시 확인할 수 있었다. 또한 전반적인 수업의 만족도에 관한 항목의 결과가 총점과 상관관계가 매우 높아서 체크리스트 10개 항목으로 구성된 강의평가의 총점이 전반적인 수업의 만족도를 충분히 반영하고 있는 것으로 보인다. 한편 체크리스트 각 항목 간의 상관관계는 높지 않아서 각각의 항목은 서로 독립적인 영역을 평가하고 있는 것으로 판단된다. 다만 '통합교육의 취지에 맞는 수업을 한다'는 항목과 '학생들의 수준에 맞는 내용이다'가 중등도의 상관관계를 보여서, 학생들의 수준에 맞는 내용으로 수업을 하는 것이 통합교육의 취지에 부합하는 것으로 인정되고 있다는 점을 확인하였다. '학습 목표를 제시한다'는 항목과 '학습 목표와 강의 내용이 일치한다'는 항목의

상관관계가 높은 것은 학습 목표를 제시하는 강의일수록 학습 목표와 내용이 일치하기 때문인 것으로 판단된다.

평가자의 주관이 개입되는 평가 결과는 평가자 간 신뢰도가 검증된 후 평가 결과가 행정적 기능을 위해 사용되어야 한다[18]. 본 연구에서는 하나의 강의에 대하여 2명의 평가자가 동시에 동료평가를 시행하도록 계획되었는데, 11개 과목의 각 강의에 대하여 두 명의 평가자 간 상관관계를 살펴본 결과, 0.968에서 0.999까지 모든 강의에서 매우 높은 수준의 상관을 보였으며, 총점 40점을 기준으로 해서 두 평가자 사이의 평균 편차는 5.03 ± 3.78 점으로 평가자 간 신뢰도가 높은 편이었다. 따라서, 본 대학에서 동료평가를 통해 교육의 수월성을 점검하고자 기대했던 방법론적으로 적절하고 평가의 양호도도 확보되었다고 할 수 있다.

한편, 학생들의 강의평가 중에서 동료평가와 비교가 가능한 문항에 대하여 상관관계를 분석하였는데, 두 문항 모두 상관관계가 낮은 것으로 나타났다. 이는 외국에서 시행된 기존의 선행 연구들과는 일치하지 않는 결과이지만[19,20], 동료평가와 학생들의 강의평가 결과가 서로 대체할 수 없는 상호보완적인 관계로서 두 평가방법 모두 교육의 수월성을 제고하는 데 의미 있는 정보를 제공해 준다고 볼 수 있다.

다만 본 연구의 분석에서 비교한 두 개의 문항이 완전히 일치한 것은 아니었으며 두 개의 문항만 비교하였기 때문에 그 결과를 일반화하는 데에는 한계가 있음에 따라 향후 추가적인 연구가 필요할 것으로 판단된다.

특히, 평가에 참여했던 평가 교수, 학생들을 대상으로 질적 인터뷰 등의 방법을 통해 교수와 학생 간 평가 결과가 일부 상이하게 나타나게 된 경향과 원인 등을 파악하기 위한 연구보완이 필요할 것으로 생각된다.

또한, 동료평가에 활용된 체크리스트 문항은 전문가 협의를 통해 개발되긴 하였으나, 평가의 수월성을 위해 10개라는 적은 수의 문항을 가지고 평가를 시행했기 때문에 풍부한 정보를 얻기가 쉽지 않은 측면이 있었으므로 향후 동료평가 체크리스트 문항이 더 신뢰할 수 있고 타당하게 구성될 수 있도록 하는 문항 개발 및 타당화 연구가 필요하겠다.

이러한 동료평가의 결과가 단순히 강의평가에 그치는 것이 아니라, 동료평가를 통해 얻은 자료를 반영하여 교육 목적과 목표를 달성하기 위한 통합교육을 설계하고 적합한 교수-

학습전략을 선택하며 통합교육의 실행과 개선 방법 등에 대한 체계적인 검토와 구체적이고 지속적인 노력을 해야 한다.

Acknowledgements: None.

Funding: None.

Conflicts of interest: None.

REFERENCES

1. Kim YI. Integrated curriculum: a pedagogical background in medical school. *Korean J Med Educ* 1998; 10: 1-10.
2. Ham TH. The approaches of the faculty to medical education at Western Reserve University. *J Med Educ* 1959; 34: 1163-1174.
3. Ahn DS, Lee YM. Current situation of medical integrated curriculum in Korea. *Korean J Med Educ* 2000; 12(Suppl): 68.
4. Kim YI, Ahn DS, Kwon BK, Kim HS, Lee BY, Cho YO. Basic study for curriculum development of medical school. Seoul, Korea: Propulsion Research Group for Medical School in Korea; 2003.
5. Lee YM, Lee KJ, Lee YJ, Ahn DS. Evaluation of integrated lectures at Korea University, College of Medicine. *Korean J Med Educ* 2000; 12: 343-351.
6. Shin JH, Hong SY. Exploration of peer review of teaching program in higher education. *J Korean Teach Educ* 2013; 30: 201-230.
7. Harris K, Farrell K, Bell M, Devlin M, James R. Peer review of teaching in Australian higher education: a handbook to support institutions in developing and embedding effective policies and practices. Carlton, Australia: The University of Melbourne, Centre for the Study of Higher Education; 2008.
8. Horowitz S, Van Eyck S, Albanese M. Successful peer review of courses: a case study. *Acad Med* 1998; 73:

- 266-271.
9. Kim S. What's new in medical education? Korean J Med Educ 1998; 10: 383-384.
 10. Lee H, Yang EB. A study on the characteristics of excellent lecturers in medical school. Korean J Med Educ 2013; 25: 47-53.
 11. Im EJ, Lee YC, Chang BH, Chung SK. Investigation of the requirements of 'good teaching' to improve teaching professionalism in medical education. Korean J Med Educ 2010; 22: 101-111.
 12. Han SI, Kim HJ, Lee JY. A comprehensive study of Korean students' evaluations of university teaching. J Educ Adm 2005; 23: 379-403.
 13. Hong KS. An analysis of students' response to course evaluation. J Educ Inf Media 2006; 12: 97-127.
 14. Woloschuk W, Coderre S, Wright B, McLaughlin K. What factors affect students' overall ratings of a course? Acad Med 2011; 86: 640-643.
 15. Oh SA, Chung EK, Rhee JA, Baik YH. An evaluation of integrated curriculum based on students' perspective. Korean J Med Educ 2007; 19: 305-311.
 16. Bacro TR, Gebregziabher M, Fitzharris TP. Evaluation of a lecture recording system in a medical curriculum. Anat Sci Educ 2010; 3: 300-308.
 17. Baek S, Im SJ, Lee SH, Kam B, Yune SJ, Lee SS, Lee JA, Lee Y, Lee SY. Instructional analysis of lecture video recordings and its application for quality improvement of medical lectures. Korean J Med Educ 2011; 23: 263-274.
 18. Seong TJ. Modern educational evaluation. 3rd ed. Seoul, Korea: Hakjisa; 2010. p 420.
 19. Wachtel HK. Student evaluation of college teaching effectiveness: a brief review. Assess Eval High Educ 1998; 23: 191-212.
 20. Kulik JA. Student ratings: validity, utility, and controversy. New Dir Inst Res 2001; 2001: 9-25.