

## Correlations of Clinical Assessment Tools with Written Examinations

Eun Jeong Choi and Sung Sunwoo

Department of Family Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

### 여러 가지 임상평가방법과 성적과의 상관관계

울산대학교 의과대학 서울아산병원 가정의학교실

최은정, 선우성

**Purpose:** Despite the goal of medical education, which is 'training to be a primary care physician,' only written examinations (WEs) generally have been used to assess medical student ability. We assessed clinical competence using the CPX (clinical performance examination), PA (portfolio assessment), VMA (video-monitoring assessment), CGD (case group discussion), and PCP (primary clinical practice) during students' clerkships in family medicine and correlated these clinical assessment tools with WEs. Also, we correlated these tools with each other to determine the feasibility of each as a replacement of the other tools.

**Methods:** Thirty-nine fourth year students at University of Ulsan College of Medicine took part in their clerkship in family medicine for 2 weeks during the first session of 2005. They took 1 CPX, 2 VMAs, 6 CGDs, and 2 PCPs and were required to submit their portfolios at the end of the clerkship. We evaluated the correlations of these tools and WEs by the overall scores of each class year and in each subject.

**Results:** The CPX, VMA, and PCP showed no significant correlation, while PA showed strong correlation with 4th year students' overall WEs. The CGD showed strong correlation with all overall scores and with almost all subjects. In addition, the PA correlated significantly with CGD.

**Conclusion:** New clinical assessment tools, such as CPX, PA, and VMA, have no significant correlation with WE, even though these tools closely correspond with real clinical practice. Therefore, these tools should be considered as complementary instruments to better assess clinical competence.

**Key Words:** Clinical competence, Documentation, Videotape recording, Clinical clerkship

Received: December 5, 2008 • Accepted: February 7, 2009

Corresponding Author: Sung Sunwoo

Department of Family Medicine, Asan Medical Center  
86 Asanbyeongwon-gil, Songpa-gu, Seoul 138-736, Korea  
TEL) 02-3010-3810 FAX) 02-3010-3815 E-mail) sws@amc.seoul.kr

Korean J Med Educ 2009 Mar; 21(1): 43-52.

doi: 10.3946/kjme.2009.21.1.43.

© The Korean Society of Medical Education.  
All rights reserved.

## 서론

의과대학의 교육목표가 일차의료에 필요한 다양한 임상능력을 갖춘 일차 의료의사의 양성이라는 점에 맞추어, 최근 들어 점차 의학적 지식의 습득 이외에도 실용적인 임상능력이 중요한 비중을 차지하게 되었다. 즉 의사의 진료 및 면담 기술, 환자를 대하는 태도 및 의사소통 능력, 질병의 예방과 건강증진 등을 포함한 임상적인 기술이 보다 중요시되면서, 이를 교육하고 평가하는 방법들이 국내 의과대학에도 도입되어 왔다. 즉 필기시험을 통한 의학지식을 평가하는 것뿐 아니라, 환자로부터 정보를 모아, 분석하고 판단하는 능력, 문제해결 능력 및 효율적인 의사소통능력 등을 포함한 복합적인 임상능력을 평가하기 위해, 객관구조화진료시험(objective structured clinical examination, OSCE), 진료수행시험(clinical performance examination, CPX), 포트폴리오 평가(portfolio assessment), 문제에 바탕을 둔 학습(problem-based learning) 등을 이용하고 있다.

이 중 실제 진료실의 상황을 재현해서 실제 임상에 가장 가까운 상황에 노출시켜 임상능력에 대한 다양한 평가를 가능하게 하는 프로그램이 바로 CPX이다. 이는 1963년 처음으로 표준화 환자의 개념이 발표된 이후[1], 현재 미국의 거의 모든 의과대학에서 사용하고 있는 가장 포괄적인 시험형태이다[2]. CPX는 캐나다, 미국의 의사자격시험에 도입되었고, 국내의 의사자격시험에도 2009년부터 도입예정이며 현재 국내 의과대학에서도 시험적 혹은 실제 평가도구로 이용되고 있다.

OSCE의 경우도 1975년 처음으로 소개된 이후[3], 국내에는 1994년 서울대학교에서 처음 도입한 이후 시험적으로 이용되고 있으며, 이에 대한 타당도 및 도입 방안 등에 대한 연구들이 진행되어 왔다[4,5,6,7,8,9,10].

CPX나 OSCE에 대한 평가의 우수성이나 타당도에 대한 연구는 이미 외국의 몇몇 연구들을 통해 입증되어 왔다. 최근 CPX나 OSCE에 대한 국내의 연구들을[4,5,6,7,11,12,13] 검토해 보면, 표준화 환자 및 교수에 의한 평가점수를 서로 비교하거나, 실습 전후의 평가점수를 비교하는 형태의 타당도 연구들이 비교적 통일된 결과를

보여주고 있다. 또한 최근에는 CPX 점수가 3학년의 OSCE성적이나 몇몇 과목의 필기시험과 연관성이 낮다는 결과를 보여주는 연구도 있었다[14].

포트폴리오 평가는 학생의 성취도 수준을 질적인 관점에서 좀 더 종합적으로 평가하기 위한 방법으로[15], 수행평가 과제나 결과물을 체계적으로 정리한 '포트폴리오(학습과제물의 모음집)'를 통해 평가하는 형태이다. 포트폴리오에는 학습자의 진료 수행이나 임상 능력을 기록한 평가지, 소그룹 활동에서의 수행평가 기록지, OSCE 시험결과, 의사소통 능력/보건의료 활동 에세이, 학생들이 임상교육에서 만났던 환자 병력청취 기록, 병원 관련 서류, 환자 질병의 진단 및 원인을 제시하는 기록, 환자들에게 명료하고 간결한 설명해주기, 처방전 기록 및 실습 기간 동안 느낀 점과 의견 등이 전체적으로 포함되므로, 그 결과분석을 토대로 자신의 장단점을 파악하여 수정, 보완할 수 있는 장점을 가지고 있다. 1993년 Royal College 의과대학과 Sheffield 의과대학에서 시험적으로 이용되기 시작하였으며, 국내에는 소수의 의과대학에서 이용하고 있는, 새로운 형태의 평가방법이다. 이러한 포트폴리오 평가는 학생의 평가뿐 아니라 지속적인 교육을 가능하게 한다는 측면에 있어서도 우수성이 점차 조명되고 있다[16,17].

비디오모니터링 평가는 학생의 병력청취 장면을 녹화하고, 이후에 비디오테이프를 다시 보면서 피드백을 통해 그룹토의를 통해 교육을 하는 방법으로, 환자 면담기법에 대한 교육을 시행하고 있는 24개 국내 의과대학 중 8개에서 사용하고 있는 교육 및 평가방법이다[18]. 이는 단순한 언어적 피드백에 비해 효과가 높다고 알려져 있으며, 전공의를 대상으로 시행한 연구를 보면 면담기술을 향상시키거나 환자 처리시간을 단축시켰다는 결과를 보여주고 있다[19,20]. 특히 지도의사의 피드백을 받는 경우에는, 자가 피드백을 받은 군이나, 대조군과 비교할 때 면담기술이나 신체진찰 기술이 모두 유의하게 호전된다는 연구도 있다[21].

학외 임상실습(의원 파견 실습교육)은 지역사회에서 실제로 일차의료를 담당하고 있는 1차 의료기관에 파견되어 실제로 현장에서 체험하면서 배우게 되는 과정이다. 이는 대학병원에서는 경험하기 힘든, 실제 일차의료

에서 흔히 다루는 문제가 무엇인지, 그리고 현장에서 어떻게 다루어지는지를 직접 체험하고 이해하는 매우 필수적인 과정이다. 현재, 가정의학 실습을 하고 있는 의과대학 25개 중 대부분인 21개의 의과대학에서 이루어지고 있으나[18], 지도의사의 교육과정이 아직 표준화되지 않았고, 역할 모델이 아직 정립되지 않은 문제점을 안고 있기도 한 평가방법이다[22,23].

이번 연구에서는 연구자들이 소속된 의과대학에서 2005년도에 가정의학과 임상실습에 참여하는 의과대학 4학년들을 대상으로 CPX, 포트폴리오 평가, 비디오 모니터링 평가, 증례 그룹토의, 학외 임상실습의 다양한 임상수행에 대한 교육과 평가를 실시하였고, 그 평가결과가 의과대학 4개년 동안의 과목별 성적 혹은 전체 성적과 어떠한 상관관계를 보이는지, 또한 평가방법들 간에는 서로 어떠한 연관성을 가지고 있는지를 알아보고, 이를 통해 전반적인 임상능력의 평가도구로 적용할 수 있는지에 대한 근거 자료를 삼고자 하였다.

## 대상 및 방법

### 1. 연구대상

2005년 3월부터 8월까지 약 6개월 동안 일개 의과대학 의학과 4학년 1학기 과정에 있는 39명의 학생을 대상으로 하였으며, 전 학생은 4~5명의 소그룹 9개조로 나뉘어, 각각 2주간의 가정의학과 실습기간 동안 다양한 교육 및 평가 프로그램에 참여하도록 하였다.

### 2. 일차의료 교육의 평가방법

#### 1) CPX증례 및 점검표 개발

본 연구에 사용된 총 9개의 시험사례에 표준화 환자 9명이 각각 동원되었다. 표준화 환자는 가정의학과 전공의 3년차가 담당하였고, 일정한 시간 동안 증례에 대해 숙지하고, 환자로서 공통된 질문사항을 암기하는 훈련과정을 거치도록 하였다. 증례 개발은 훈련된 가정의학과 전문의에 의해 일차의료에서 흔히 접할 수 있는 의학적 문제 9개의 증례를 선별하여, '서울CPX컨소시엄'에서 개

발한 CPX의 증례들과 동일한 형식과 유형에 따라 개발하였고, 가정의학과 교수에 의해 점검 및 보완되었다. 점검표의 문항 수는 18~19문제로 하였고, 병력청취에 대한 평가문항들(20점), 질환에 대한 설명 및 환자교육에 대한 평가문항(15점), 진료 전반에 대한 만족도(5점), 의사의 임상예절 및 의사-환자의 관계(5점), 재방문에 대한 환자로서의 의지여부(5점) 등의 평가문항들을 포함하였다.

CPX의 시행은 실습 2주차 마지막 날에 진료실을 대상으로 만들어진 모의 진료실에서 표준화 환자(가정의학과 전공의 3년차)를 이용하여 외래 환자 진료의 능력 및 태도를 평가 받도록 하였고, 시험과정을 모두 녹화하여 필요할 때는 다시 점검할 수 있도록 하였다. 시험에 대한 시간배정은 학생별로 모두 12분으로 하였으며, 채점에는 표준화 환자(가정의학과 전공의 3년차), 가정의학과 교수 1인 및 전문의 1인, 가정의학과 전공의 3년차 1인으로 구성된 총 4명이 팀이 되어 참여하였고, 병력청취 중점으로 평가하였다.

#### 2) 포트폴리오 평가

가정의학과 실습기간 동안에 첫째 주 시작에서부터 둘째 주 실습이 끝날 때까지의 전 일정에 대한 과제로, 학생이 실습과정 중에 준비한 과제물 및 예습, 복습한 내용, 검토한 문헌, 실제 자신이 담당했던 환자의 병력 청취기록지, 증례 그룹토의 준비 내용, 실습에 대한 관심 및 참여도, 학외실습 감상문, 실습 후 feedback을 위한 설문지를 첨부하도록 하였다. 상기 자료들을 검토하여 채점항목에 다음 10가지 즉 조별 평가, 전반적인 수행도, 성실성, 협동성, 적극성, 비디오 모니터링에 대한 소감, 학외실습에 대한 감상문, 실습 전반을 통해 느낀 점, 예습, 복습 등을 포함하였다. 실습기간이 끝나는 날 포트폴리오를 제출하였고, 100점 만점으로 실습과정을 주도한 가정의학과 교수가 직접 채점하였다.

#### 3) 비디오 모니터링 평가

가정의학과 외래를 방문한 초진환자를 대상으로 하여, 2주 동안의 가정의학과 실습 기간 동안 1주에 최소 2~3회 이상 예진을 시행하였다. 먼저 환자의 동의를 구한 뒤 동의한 경우에 한하여, 실제 진료실에서 환자의 병력청취를 시행하고, 병력기록지를 작성하였다. 그 중에서 적

합한 증례를 선택하여 학생 1인당 매주 1회씩 총2회 비디오로 녹화하였다. 병력청취는 환자의 주증상에 따라 대개 15~20분의 시간을 엄수하도록 하였고, 이미 만들어진 점검표의 세부 평가항목 마다 1~5점으로 20개 항목을 평가하였다. 훈련된 가정의학과 전문의에 의해 채점하되 두 증례 중 더 나은 점수를 선택하였고, 평가 시에는 평가자와 소그룹의 학생들이 함께 비디오를 청취 및 분석하면서 병력청취 중 부족했던 점이나 배울 점, 비언어적 의사소통 등에 대해 의견을 공유하는 과정을 통해 교육하고 평가하였다.

#### 4) 증례 그룹토의 및 평가

2주간의 일정동안 각 그룹별로 총 6회의 정해진 주제에 따라 증례 그룹토의가 시행되었고 가정의학과 교수 및 전임의 6인이 각각의 증례를 지도하였으며, 시간은 60~90분 동안 시행되었으며, 토의가 끝나면 담당했던 교수진에 의해서 토의의 준비도 및 참여도가 평가되었다. 미리 증례와 참고문헌을 지정해 주어 얼마나 철저히 준비하여 토의에 임하였는지 여부, 증례 및 문헌내용에 대한 이해도 등 지식에 대해 50%, 발표의 적극적인 참여도를 50%로 하여 채점하였다.

#### 5) 학외 임상실습(의원 파견 실습)

지역사회에서 개원하여 1차 의료를 실제로 행하고 있는 각각 다른 현장에 2일간 파견되어 일차 진료에 관한 지식이 실제 진료에서 어떻게 적용되고 있는가를 학습하고, 참관한 진료내용을 정리하며 진료내용에서 얻은 질문 및 학습에 대한 보고서를 과제로 제출하도록 하였다. 증례보고 및 개원의가 평가한 자료를 참고로 외래 환자 진료의 상담 및 임상지식 50%, 진료 태도 50%로 평가하였다.

#### 6) 출석

학교에서 권장하는 기준인 전체점수의 10%에 준하여 전체 200점 중 20점을 출석점으로 배정하였고, 무단결석 시 5점을 감점하고 지각 및 조퇴 시 1점을 감점하여 점수를 산정하였다.

#### 7) 실습 배점 비율

CPX	포트폴리오 평가	비디오 모니터링 평가	증례 그룹토의	학외 임상실습	출석
30	100	10	20	20	20

가정의학과 실습 점수는 실습기간 중에 차지하는 시간 및 중요도의 비중을 고려하여, 임상능력을 평가하는 지표가 총 200점 중 100점(CPX+포트폴리오 중 60점+비디오모니터링)으로 50%를, 의학지식에 대한 비중은 200점 중 50점(포트폴리오 30점+증례그룹토의 20점)으로 25%를, 그 밖에 학외실습 10%, 출석 10%, 전반적인 학습태도 5%(포트폴리오 중 10점)의 비중으로 배점하였다.

### 3. 학업성취도

의학과 4개 학년 동안에 치루는 필기시험과 실습성적을 모두 포함하였으며, 전체총점뿐만 아니라 각 학년별 총점, 전 교과과정의 과목별 성적을 포함하였다.

### 4. 분석

각 점수들의 관련성은 Pearson상관계수로, 일부 비정규 분포를 보이거나 순위척도로 채점된 항목에 대해서는 Spearman 상관계수를 구하였다. 상관계수가 -0.1~0.1 사이이면 상관관계를 무시할 수 있으며, 0.1과 0.3 사이면 약한 양적 선형관계, 0.3과 0.7 사이면 뚜렷한 양적 선형관계, 0.7과 1.0 사이면 강한 양적 선형관계로 간주하였다. 양측검정을 통해 통계기법에 따라  $p < 0.01$  또는  $p < 0.05$ 를 의미 있는 것으로 인정하였으며, 통계 프로그램은 SPSS version 12.0 (SPSS Inc, Chicago, IL, USA)을 사용하였다.

## 결과

### 1. 연구대상자의 성적분포

가정의학과 실습에 참여한 의과대학 4학년 학생은 39명이었고, 참가자 전원이 4~5명의 소그룹별로 2주간의 실습기간 동안 모든 평가방법을 완전히 수행하였고 참가자의 성별 분포는 남자 24명, 여자 15명이었다. 학생들의 전 학년 평점은 4.5점 만점을 기준으로 평균 3.3점, 표준오차 0.5점으로 정규분포를 하였으며, 가정의학과 실습성적은 200점 만점을 기준으로 평균 180.0점을 보였다. CPX, 비디오 모니터링 평가, 증례그룹 토의, 포트폴

Table 1. Score Distribution of All Participants (n=39)

	Full marks	Mean	Min	Max	SD
Overall score	4.5	3.3	2.1	4.3	0.5
Family medicine clerkship	200	180.0	149.7	185.9	0.8
CPX	30	16.6	9.5	23.4	2.9
Video monitoring	10	9.1	7.7	9.9	0.5
Case group discussion	20	17.8	12.0	20.0	1.2
Portfolio assessment	100	95.6	79	99	3.5

CPX: Clinical performance examination, SD: Standard deviation.

Table 2. Correlation of Clinical Assessment Tools with the Overall Scores in Each Class Year (n=39)

Grade	CPX	Portfolio assessment	Video monitoring	Case group discussion	Primary care practice
Overall	0.111	0.242	0.010	0.517 <sup>bl</sup>	0.269
Fourth	0.089	0.369 <sup>al</sup>	0.093	0.619 <sup>bl</sup>	0.243
Third	0.112	0.305	-0.058	0.451 <sup>bl</sup>	0.269
Second	0.095	0.256	0.003	0.533 <sup>bl</sup>	0.220
First	0.098	0.049	0.043	0.378 <sup>al</sup>	0.222

CPX: Clinical performance examination.

<sup>a)</sup>p<0.01, <sup>b)</sup>p<0.05.

리오 평가점수는 극단값을 제외하고는 평균을 중심으로 정규분포를 보였다(Table 1).

## 2. 학년별 및 전 학년 평점과 여러 임상평가 방법들 간의 상관관계

CPX성적과 각 학년별 평점 및 전체 학년의 평점과의 상관관계는 통계적으로 유의하지 않았으며 상관계수를 보면 0.10 내외로 매우 낮았다. 비디오 모니터링 평가나 학외 임상실습의 경우도 기존 성적과는 유의한 상관관계를 보이지 않았다. 포트폴리오 평가의 경우에는 임상실습이 이루어지는 4학년 평점과 유의한 상관관계를 보였으며( $r=0.369$ ,  $p<0.01$ ), 저학년보다 고학년이 될수록 상관계수가 높아지는 경향을 보였다(Table 2).

의학지식을 위주로 평가되는 증례 그룹토의는 기존 성적과 가장 뚜렷한 양적 상관관계를 보였으며 모든 학년의 학년별 평점 및 전체평점에서 모두 유의한 연관성을

보였다( $p<0.05$ )(Table 2).

## 3. 과목별 성적과 평가방법들 간의 상관관계

의과대학 4년간의 전 교과과정 중에 있는 모든 과목들과 임상평가 방법들 간의 상관관계를 분석하였다. 전 과목을 포함하였으나 4학년 임상실습 중 선택과목인 비주요과목들은 필수과정에 포함되지도 않을 뿐 아니라 학생마다 선택한 항목이 모두 다르고 학점의 비중도 매우 적어 분석에서 제외하였다(Table 3).

CPX점수는 각 과목별 점수와 통계적으로 유의한 상관관계를 보이지 않았다. 하지만 과목별 경향을 보면 다른 과목에 비해 내과 과목, 특히 내과실습성적에서 다른 과목에 비해 상관계수가 높았다( $r=0.213$ ). 비디오 모니터링 평가나 포트폴리오 평가의 경우에도 특정 과목과 상관관계를 보이지 않았다(Table 3).

증례 그룹토의의 경우는 거의 모든 과목성적과 뚜렷한

Table 3. Correlation of Clinical Assessment Tools with Scores in Each Subject for the Entire School Period (n=39)

Name of subject	CPX	Portfolio assessment	Video monitoring	Case group discussion
Introduction of medicine	-0.001	-0.117	-0.029	0.341 <sup>a)</sup>
Infection	0.214	-0.170	0.189	0.428 <sup>b)</sup>
Nephrology	-0.010	0.017	0.184	0.496 <sup>b)</sup>
Cardiovascular	0.196	-0.047	-0.103	0.273
Immunology	0.149	-0.049	-0.225	0.294
Pulmonology	0.116	-0.004	-0.017	0.430 <sup>a)</sup>
Gastroenterology	0.055	-0.037	-0.048	0.365 <sup>a)</sup>
Hemato-oncology	0.139	-0.222	-0.084	0.223
Growth and aging	0.056	0.095	0.179	0.376 <sup>a)</sup>
Endocrinology	0.053	-0.069	0.014	0.455 <sup>b)</sup>
Reproductive medicine	-0.080	0.090	0.011	0.294
Myology/osteology	-0.044	0.047	-0.096	0.379 <sup>a)</sup>
Neurology	0.080	0.006	-0.033	0.412 <sup>b)</sup>
Psychiatry	0.161	0.038	-0.095	0.392 <sup>a)</sup>
Family & community medicine	0.143	0.073	0.122	0.429 <sup>b)</sup>
Preventive medicine	0.184	-0.022	-0.244	0.108
Medical history and ethics	0.137	-0.073	-0.143	0.265
Introduction of clinical medicine	-0.095	0.141	0.106	0.202
Integrated medicine	0.177	0.235	0.122	0.404 <sup>a)</sup>
Psychiatry clerkship	-0.006	0.025	-0.173	0.175
Neurology clerkship	-0.142	0.150	0.007	0.290
Internal medicine clerkship	0.213	0.041	-0.118	0.255
Pediatrics clerkship	0.123	0.145	0.083	0.118
Obstetrics/Gynecology clerkship	0.139	0.020	-0.157	0.034
Integrated examination	0.067	0.058	0.055	0.226
General surgery clerkship	0.165	0.133	-0.268	0.038
Clinical lecture 1	-0.029	0.077	0.028	0.365 <sup>a)</sup>
Clinical lecture 2	-0.029	0.059	-0.001	0.212
Clinical lecture 3	-0.099	0.039	-0.048	0.299
Emergency medicine clerkship	0.063	-0.100	-0.024	-0.150
Clinical lecture 4	0.099	0.073	0.069	0.253
Clinical lecture 5	-0.208	-0.073	0.131	0.249
Health management	0.132	0.010	-0.039	0.404 <sup>a)</sup>
Integrated examination 1	0.147	0.138	0.137	0.336 <sup>a)</sup>
Integrated examination 2	0.019	0.189	0.071	0.348 <sup>a)</sup>
Integrated examination 3	-0.001	0.080	0.051	0.374 <sup>a)</sup>

CPX: Clinical performance examination.

<sup>a)</sup>p<0.01, <sup>b)</sup>p<0.05.

상관관계를 보였다. 3학년 교과과정인 외과 임상실습(r=0.038), 산부인과 임상실습(r=0.034)에서는 상관관계를 보이지 않았고, 응급의학실습(r=-0.150), 예방의학(r=0.108), 정신과 임상실습(r=0.175)과 소아과 임상실습(r=0.118)에서도 낮은 상관계수를 보였으나, 그 외의

거의 모든 과목에서 통계적으로 유의한 뚜렷한 상관관계를 보였다. 즉 증례 그룹토의의 경우 주요 과목들의 임상실습 점수와는 상관관계를 보이지 않았던 반면, 오히려 필기시험으로 100% 평가되는 과목들이나 3회의 종합평가시험(r=0.336, r=0.348, r=0.374)에서 연관성이 더욱



Table 4. Correlation of Clinical Assessment Tools with Each Other (n=39)

	Video monitoring	Case group discussion	Portfolio assessment	Primary care practice
CPX	0.035	0.135	0.265	-0.042
Video monitoring		0.159	0.287	-0.108
Case group discussion			0.699 <sup>a)</sup>	0.088
Portfolio assessment				-0.073

CPX: Clinical performance examination.

<sup>a)</sup>p<0.01.

뚜렷하게 나타났다(p<0.05)(Table 3).

#### 4. 임상평가 방법들 간의 상관관계

CPX의 경우 포트폴리오 평가와의 상관계수는 0.265 였으나 통계적으로 유의한 수준은 아니었고, 비디오 모니터링 평가, 증례 그룹토의 및 학외 임상실습 점수와는 뚜렷한 상관관계를 보이지 않았다. 비디오 모니터링 평가 또한 포트폴리오 평가와 0.287로 가장 높은 상관계수를 보였으나 통계적 유의성은 없었고, 다른 평가방법들과도 상관관계를 보이지 않았다. 증례 그룹토의의 경우는 포트폴리오 평가점수와 통계적으로 매우 뚜렷한 상관관계를 보였으며(r=0.699), 학외 임상실습과 다른 평가 방법 간에 연관성은 관찰할 수 없었다(Table 4).

## 고찰

본 연구를 통하여 기존의 평가도구인 시험성적과 CPX, 포트폴리오 평가, 비디오 모니터링 평가 등 임상평가방법들 간의 상관관계 및 평가방법들 간의 상호 연관성을 알아봄으로써, 기존의 성적 평가결과와 차이점을 보이는 지에 대해 살펴보고, 이러한 평가방법들을 전반적인 임상실습 평가 도구로 적용할 수 있는지에 대한 가능성을 가늠해 보고자 하였다.

이번 연구에서 CPX와 성적간의 상관관계를 보면 학년별 평점이나 전 학년의 평점과의 상관계수가 0.11로, 이전 연구에서 보였던 0.28~0.32[13,14]에 비해 낮은 수치를 보였는데 그 이유로 본 연구에서 CPX 시험 내용에

신체검진을 포함시키지 않았던 점을 이유로 들 수 있다. 이전의 연구에 따르면 CPX의 신체검진 항목은 학교성적과 연관성이 높은 반면에 병력청취의 경우 연관성이 없거나 상관관계가 약하다는 동일한 결과가 있었는데 [13,14], 본 연구에서 신체검진을 포함시키지 않았던 점을 고려하면 상관계수가 더 낮게 나타난 것이 당연한 결과라고 볼 수 있다. 즉 신체검진보다 병력청취는 학교성적과의 연관성이 더 적고, 의학적인 지식 습득만으로 얻을 수 없는, 즉 임상경험이나 훈련 및 교육과정을 통해서 얻어질 수 있을 것이라는 점을 알 수 있다.

또한 이번 연구 결과 CPX점수는 기존의 성적과 상관관계를 보이지 않았을 뿐 아니라 특정 과목과의 상관관계도 보이지 않았다. 하지만 과목별로 볼 때 내과과목이나 내과임상실습 성적이 다른 과목보다는 비교적 높은 상관계수(r=0.21)를 보이고 있다. 이는 이전의 연구에서 [14] 보인 결과와 일치하는 것으로 CPX의 증례가 대개 내과적 질환을 중심으로 하고 있고 병력청취에 대한 평가항목이 내과적 지식에 좀 더 치우쳐 있기 때문이라고 생각할 수 있다.

다른 임상평가 방법들과의 연관성 분석에서 CPX는 포트폴리오 평가, 비디오 모니터링 평가, 증례 그룹토의 혹은 학외 임상실습 점수와 상관관계를 보이지 않았다. 병력청취를 위주로 하는 비디오 모니터링 평가와의 공통점을 고려하면 연관성을 기대해 볼 수 있지만, 비디오 모니터링 평가에서는 주로 병력청취의 내용에 초점을 맞추었던 반면에, CPX 경우는 실제 임상능력 중에 매우 중요한 부분인 환자의 만족도나 의사의 예절 및 환자-의사 관계, 재방문의 의지 혹은 환자교육에 대한 평가점수가

60%를 차지하였던 점이 다르기 때문에 두 가지 평가방법 간에 상관관계를 보이지 않았을 것이라고 분석해 볼 수 있다.

이번 연구에서 시도했던 포트폴리오 평가의 경우는 기존 성적과의 연관성에 대한 분석에서 4학년의 성적과 유의한 연관성을 보이며, 1학년 성적을 제외한 모든 학년별 및 전 학년 평점과 0.242~0.369의 상관계수를 보이고 이는 고학년으로 갈수록 높아지는 경향을 보였다. 또한 다른 임상평가 방법과의 분석에서 증례 그룹토의 점수와 통계적으로 뚜렷한 상관관계( $r=0.699$ )를 보였고, 비디오 모니터링( $r=0.287$ )이나 학외실습 점수( $r=-0.073$ ), CPX 점수( $r=0.265$ )와는 통계적으로 유의한 연관성은 보이지 않았다.

물론 포트폴리오 평가에 대한 연구가 국외나 국내 모두 매우 부족한 상태이므로 아직 어떠한 결론을 내리기에 이르지만, 포트폴리오 평가 자체의 포괄성을 감안한다면, 기존 성적뿐만 아니라 다른 평가방법들과 적절한 연관성을 보일 수 있을 것이라고 예측해 볼 수 있고, 부족하나마 이번 연구에서 그러한 예측을 뒷받침해주고 있다. 물론 아직 정형화된 포트폴리오의 평가기준이 확립되지 않은 점, 실제 실습 일정이 2주 동안이면서 그 기간 중 바쁜 실습 일정을 따라야 하므로 부족했던 준비 기간, 혹은 실습에 참여한 시점이 학생마다 다르기 때문에 발생하는 준비 기간의 차이, 포트폴리오에 대한 정보 공유로 인해 발생하는 준비 요령의 차이 등을 예상해 볼 때, 그 연관성을 더욱 민감하게 밝힐 수 있도록 하기 위해서는 이러한 점들을 고려하고 통제하는 과정이 필요할 것으로 사료된다. 하지만 이번 연구에서 얻은 결과를 토대로 볼 때 포트폴리오 평가는 임상평가 도구들 중에 가장 포괄적으로 접근할 수 있는 방법이면서, 여러 평가방법에 따른 성적이 심한 변이를 가진 경우에 이를 보정하는 역할을 해줄 수 있는 새로운 형태의 평가방법으로 주목받을 수 있겠다는 의미를 부여할 수 있다.

비디오 모니터링 평가는 기존의 학년별 및 과목별 성적과 상관관계를 보이지 않았으며, 학외 임상실습 또한 학년별 성적과의 연관성이 없었다. 또한 이 두 가지 임상평가 방법 모두 다른 평가방법과의 연관성을 보이지 않았다.

증례 그룹토의는 기존의 모든 학년별 성적 및 4년간의 총 평점과 매우 뚜렷한 상관관계를 보였으며, 과목별 성적에서도 거의 대부분의 과목들 성적과 유의한 상관관계를 보였다. 즉 3학년, 4학년의 교과과정 중에 있는 주요 과목 임상실습을 제외한 거의 모든 과목에서 뚜렷한 상관관계를 보였으며, 특히 1,2학년의 교과과정에 있는 과목들이나, 3,4학년의 종합평가시험과 같은 필기시험으로 평가되는 과목에서 상관계수가 더욱 높게 나타났다. 이러한 결과는 증례 그룹토의의 과정 자체가 미리 선택된 증례에 대한 연구 및 참고문헌을 검토하는 연습을 통해 준비하고 나서 토의내용을 평가하는 것이었으므로 의학 지식에 대한 평가인 필기시험 성적과 뚜렷한 상관관계를 보이지만, 반면에 병원에서 임상실습 과정으로 평가되는 과목에 대해서는 낮은 상관관계를 보이게 된 것으로 생각해 볼 수 있다.

이전의 연구들과 달리 이번 연구에서는 임상평가 방법의 성적들을 의과대학 교육과정 전체 과목의 성적과 비교하는 것이 가능하였는데, 그 이유는 CPX 시험을 포함한 가정의학과 실습이 시행된 시기가 주요과목의 임상실습을 모두 마친 4학년 1학기였고, 실습이 시행된 의과대학에서 가정의학과 실습이 필수과목으로 지정되어 있었기 때문에 가능했다.

물론 가정의학과 실습 시기가 3월에서 8월 사이로 학생들마다 같지 않았던 점이 비뚤림(bias)으로 작용했을 수 있지만, 차이가 나는 실습기간 동안에는 학생들이 비주요과목의 선택실습과정 중에 있었기 때문에 교과과정의 차이 즉 특정 주요과목에 대한 지식의 차이로 생기는, CPX나 포트폴리오 평가의 차이는 거의 미미할 것으로 예상할 수 있다. 또한 실습시기가 다른 소그룹마다 증례에 대한 정보를 공유하여 CPX 성적의 차이가 초래되는 것을 막기 위해 매년 다른 증례를 개발하여 적용하였고, 시험 직전까지 증례에 대한 보안이 유지되었으며, 표준화 환자를 외래환자에 대한 임상경험이 풍부하고, 적절하게 훈련받은 가정의학과 전공의 3년차로 하였으며, 표준화 환자와 교수를 포함한 가정의학과 의사 4인이 팀이 되어 평가에 참여하였기 때문에 한 사람의 표준화 환자에 의한 평가에 비해 주관적으로 평가되는 것도 최대한 보정되었으리라고 생각한다.



이번 연구는 한 학년만을 모집단으로 평가하였으며, 학생수도 소규모이므로, 개인별 수학적능력이나 학년의 분위기에 따른 편차가 발생할 수 있다는 제한점이 있다. 그렇기 때문에 같은 방법을 이용하여 3~5년 이상 연구데이터를 축적한 뒤 분석한다면, 향후 좀 더 객관성 있는 자료가 마련될 수 있을 것이다.

현재까지 의과대학에서 의학교육의 목표를 달성하기 위해 일차의료인으로서 필수적으로 갖추어야 할 의학적 지식 이외의 임상능력들을 기존의 평가방법만으로 평가하기가 힘든 사실로 인해서, CPX를 비롯한 새로운 평가방법들에 대한 연구가 점차 활발하게 이루어지고 있고, 그 우수성 및 타당성이 점차 밝혀지고 있다. 이번 연구를 통해 CPX, 비디오모니터링 평가, 포트폴리오 평가, 증례 그룹토의, 학외 임상실습을 이용한 실습평가를 시행한 결과, 증례 그룹토의를 제외한 평가방법들의 점수가 기존의 필기시험 성적과 뚜렷한 상관관계가 없다는 결과를 얻을 수 있었다. CPX 등의 임상평가 도구들이 임상능력을 평가하는 우수성이 입증되고 있는 이상 이러한 결과를 토대로 기존의 성적평가 방법만으로는 임상능력을 충분히 평가해내기가 부족하리라는 결론을 얻을 수 있다. 앞으로 보다 나은 일차의료인을 양성해 내려는 의학 교육의 발전을 위해 이러한 부족한 부분을 보완하려면, 앞서 이용한 새로운 평가도구가 좀 더 표준화되고 실용화되어야 할 것이며, 그러기 위해서는 앞으로도 많은 노력과 연구과정이 필요할 것이다.

## REFERENCES

1. Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. *J Med Educ* 1964; 39: 802-805.
2. Lee HS. Primary care education in medical schools in USA. *J Korean Acad Fam Med* 1999; 20: 761-770.
3. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J* 1975; 1: 447-451.
4. Jung KH, Choi EH, Bae JD, Na DY, Son IK, Ko CW, et al. Comparison of professional standardized and amateur standardized patients in OSCE. *Korean J Med Educ* 2005; 17: 97-105.
5. Park H, Han J, Park M, Oh J. Comparison of results from objective structured clinical examinations for medical students performed before and after clinical clerkship. *Korean J Med Educ* 2004; 16: 63-71.
6. Han JJ, Park H, Eo E, Yoo K, Lee D, Jung WS. An OSCE for summative assessment after clinical clerkship: experience in Ewha Medical School. *Korean J Med Educ* 2004; 16: 33-40.
7. Park H, Lee J, Hwang H, Lee J, Choi Y, Kim H, et al. The agreement of checklist recordings between faculties and standardized patients in an objective structured clinical examination (OSCE). *Korean J Med Educ* 2003; 15: 143-152.
8. Park H. Clinical application of objective structured clinical examination (OSCE). *Korean J Med Educ* 2004; 16: 13-23.
9. Kim SH, Lee DH, Lee YM, Choi YS, Cho KH, Hong MH. Assessing the clinical competence of family medicine clerkship students using the OSCE. *J Korean Acad Fam Med* 2002; 23: 583-592.
10. Lee YM, Hwang K, Baik SH, Hong MH, Choi YS, Kim SH. An experience of OSCE for introducing clinical performance assessment to Korean medical licensure examination. *Korean J Med Educ* 2001; 13: 47-57.
11. Kwon I, Kim N, Lee SN, Eo E, Park H, Lee DH, et al. Comparison of the evaluation results of faculty with those of standardized patients in a clinical performance examination experience. *Korean J Med Educ* 2005; 17: 173-184.
12. Shin J, Lee S, Park H. Standardized patients' accuracy in recording checklist items during clinical performance examinations. *Korean J Med Educ* 2005; 17: 197-203.

13. Kim JJ, Lee KJ, Choi KY, Lee DW. Analysis of the evaluation for clinical performance examination using standardized patients in one medical school. *Korean J Med Educ* 2004; 16: 51-61.
14. Park WB, Lee SA, Kim EA, Kim YS, Kim SW, Shin JS, et al. Correlation of CPX scores with the scores of the clinical clerkship assessments and written examinations. *Korean J Med Educ* 2005; 17: 297-303.
15. Kim S. Portfolio assessment and its application in medical education. *Korean J Med Educ* 2003; 15: 73-81.
16. Webb C, Endacott R, Gray MA, Jasper MA, McMullan M, Scholes J. Evaluating portfolio assessment systems: what are the appropriate criteria? *Nurse Educ Today* 2003; 23: 600-609.
17. Finlay IG, Maughan TS, Webster DJ. A randomized controlled study of portfolio learning in undergraduate cancer education. *Med Educ* 1998; 32: 172-176.
18. Choi CJ, Kim BT, Ou SW, Won CW, Yoo SM, Park H. The survey of family medicine clinical clerkships in medical schools in Korea. *J Korean Acad Fam Med* 2004; 25: 895-900.
19. Scherer LA, Chang MC, Meredith JW, Battistella FD. Videotape review leads to rapid and sustained learning. *Am J Surg* 2003; 185: 516-520.
20. Chou C, Lee K. Improving resident's interviewing skills by group videotape review. *Acad Med* 2002; 77: 744.
21. Scheidt PC, Lazowitz S, Ebbeling WL, Figelman AR, Moessner HF, Singer JE. Evaluation of system providing feedback to students on videotaped patient encounters. *J Med Educ* 1986; 61: 585-590.
22. Kim BT, Shin DH. Ui-won pagyeon silseub gyo-yug. *J Korean Acad Fam Med* 2004; 25: S394-S398.
23. Lee JK, Park HK. Evaluation of community primary care clerkship. *J Korean Acad Fam Med* 1999; 20: 812-821.