

Content Analysis of Standardized-Patients' Descriptive Feedback on Student Performance on the CPX

Young Hee Lee, Young-Mee Lee and Byung Soo Kim

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

의과대학생의 임상수행능력에 대한 표준화 환자의 서술형 평가내용 분석

고려대학교 의과대학 의학교육학교실

이영희, 이영미, 김병수

Purpose: The goal of this study was to explore what kind of additional information is provided by the descriptive comments other than the rating scales, on the physician-patient interaction (PPI) in the clinical performance examination (CPX) and its feedback role in identifying students' strengths and weaknesses in communication skills.

Methods: The data were collected from 18 medical schools in Seoul and Gyeonggi region, which participated in the CPX for fourth-year medical students in 2006 and 2007. In total 12,650 examination cases in 2006 and 12,814 cases in 2007 were analyzed. Descriptive comments from the standardized patients (SPs) were analyzed by content analysis, which includes a 4-step process: coding, conceptualizing, categorizing and explanation.

Results: Ten categories (41 concepts) for 'strength' and 11 for 'weakness' (40 concepts) in the PPI were extracted. Among them, 10 categories were the same in both strength and weakness: providing adequate interview atmosphere, attentive listening, providing emotional support, non-verbal behaviors, professional attitude, questioning, explanation, reaching agreement, counseling & education and conducting adequate physical examination. For the 'structured and organized interview', only weakness was described. In 'providing emotional support' and 'adequate interview atmosphere', comments on strengths were more frequently mentioned than weaknesses. However, communication skills that were related to non-verbal behaviors were more frequently considered weaknesses rather than strengths. The numbers and content of the SP's comments on students' strengths and weaknesses in the PPI varied depending on the case specificities.

Conclusion: The results suggest that the SPs' descriptive comments on student' performance on the CPX can provide additional information versus structured quantitative assessment tools such as performance checklists and rating scales. In particular, this information can be used as valuable feedback to identify the advantages and disadvantages of the PPI and to enhance students' communication skills.

Key Words: Physician-patient relations, Clinical competence, Qualitative research

Received: October 8, 2010 • Revised: November 5, 2010 • Accepted: November 3, 2010

Corresponding Author: Young-Mee Lee

Department of Medical Education, Korea University College of Medicine, 126-1 Anam-dong 5ga, Seongbuk-gu, Seoul 136-705, Korea

Tel: +82.2.920.6098 Fax: +82.2.928.1647 email: ymleehj@korea.ac.kr

Korean J Med Educ 2010 Dec; 22(4): 291-301.

doi: 10.3946/kjme.2010.22.4.291.

pISSN: 2005-727X 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

학생의사의 임상수행능력을 평가하는 유용한 방법 중의 하나는 표준화 환자(standardized patients)를 활용하는 것이다. 표준화 환자는 신체진찰 기술을 가르치거나[1], 커뮤니케이션 기술을 평가하거나[2], 나쁜 소식 전하기와 어려운 환자 다루기 같은 상황에 학생들을 노출시킬 때 사용되어왔다[3,4]. 표준화 환자는 의도적으로 설계된 의료면담 상황에서 실제 환자의 입장에서 학생들의 다양한 행동과 태도에 대해 정확한 평가와 형성적 피드백을 제공할 수 있는 장점이 있다[5]. 2006년부터 현재까지 수행되어온 임상수행능력시험(clinical performance examination, CPX) 평가양식은 구조화된 양적 평가와 서술형 평가로 구성되어 있다. 구조화된 평가양식은 모든 사례에 공통적인 항목으로서 진료 전반에 대한 만족도, 인사 및 소개, 병력청취, 환자-의사관계를 평가하도록 구성되어 있고, 사례에 따라 병력청취, 정보나누기, 신체진찰, 환자교육, 의사예절 등을 포함하고 있다. 서술형 평가양식에는 표준화 환자가 학생의사에게 꼭 해주고 싶은 말을 학생의사가 잘한 점과 앞으로 고쳐야 할 점으로 나누어 자유롭게 기술하도록 하고 있다.

구조화된 평가표의 환자-의사관계는 7개 항목으로 구성된 평점표에 따라 포괄적으로 평가하도록 되어 있기 때문에 환자-의사 관계에서 학생의사가 수행해야 할 핵심 커뮤니케이션 행동의 수행 여부를 세밀하게 평가하기 어렵다. 반면에 서술형 평가는 구조화된 항목으로 평가하지 못한 유의미한 내용을 CPX 상황에 따라 생생하게 기록하므로 개별적이고 다양하며 심층적인 정보를 제공해 준다. 또한 서술형 평가내용은 표준화 환자가 학생들의 행동과 태도 중 인상 깊었던 것을 2~3분 이내의 짧은 시간동안 작성하기 때문에 면담기간동안 가장 크게 각인된 인상과 느낌만이 기록된다. 즉, 서술형 평가의 환자-의사 관계 내용은 7개 항목으로 묶어 포괄적으로 평가하고 있는 구조화된 환자-의사 관계 채점표가 놓칠 수 있는 개별적이고 세부적인 정보를 제공함으로써 구조화된 평점표의 기능을 보완할 수 있다. 그 동안 CPX 관련 선행 연구들은 구조화된 평가지의 양적 데이터를 활용한 연구들이 주를 이루었고[6,7,8], 서술형 평가내용을 대상으로 한 연구는 없었다.

연구자들은 CPX 평가체제에서 서술형 평가의 기능을 점검해 보고, 그것의 교육 평가적 의미와 활용가능성을 검토해 보고자 하였다. 즉, CPX에서 표준화 환자가 학생의사의 수행에 대해 기술한 자유기술형 피드백 내용을 체계적으로 분석하여 환자-의사 관계 측면에서 학생의사의 '강점'과 '약점'을 파악하고, 이러한 자유기술형 피드백이 정량적 평점표에 비하여 환자-의사 관계 영역의 수행능력에 대해 추가적으로 제공하는 정보가 무엇인지를 파악하는 것이 이 연구의 목적이다.

대상 및 방법

1. 연구 대상 및 자료 수집

서울경기 CPX 컨소시엄 소속 대학 중 2006년 및 2007년에 CPX 시험에 참여하였던 18개 대학을 연구대상으로 하였다. 시험에 응시했던 학생 수는 2006년 1,732명, 2007년 1,681명이었다. 각 대학에서 시험을 치렀던 CPX 문제 수는 학생 1인당 최소 5개부터 최대 10개까지였으며, 총 분석 사례 수는 2006년 12,650개, 2007년 12,814개였다(Table 1). 표준화 환자의 서술형 평가 내용 중 분석범위는 '학생의사가 잘한 점'과 '학생의사가 앞으로 고쳐야 할 점'이라고 기록한 내용 중 환자-의사 관계부분에 한정하였으며, 모든 학생에게 동일한 피드백을 제공하여 신뢰성이 낮다고 판단된 서술형 기록은 분석대상에서 제외하였다.

2. 연구 방법 및 절차

1) 연구 방법

이 연구의 분석대상이 된 서술형 평가 자료는 이론적 배경을 토대로 수집된 자료라기보다 CPX 현장에서 학생의사의 행동에 대한 표준화 환자의 생각과 느낌을 구조화된 틀에 얽매이지 않고 직접 기록하도록 한 실제적 자료였다. 기록물로서의 자료의 특성에 맞게 이 연구는 질적 연구 방법과 연구 절차를 따랐다. 질적 연구는 텍스트(text)와 이미지(image) 자료에 의존하고, 자료 분석에 있어서 독특한 단계(step)를 가지며, 다양한 탐구전략을 선택한다는 점에서 양적 연구와 다르다[9,10]. 선택된 탐구전략은 연구 절차에 지대한 영향을

Table 1. The Subjects of the Study

College of Medicine ^{a)} (n = 18)	2006			2007		
	Students (A) ^{b)}	Test cases per student (B)	Analyzed cases (A*B)	Students (A) ^{a)}	Test cases per student (B)	Analyzed cases (A*B)
A	38	6	228	42	10	420
B	58	8	464	69	8	552
C	119	6	714	125	6	750
D	131	5	655	134	6	804
E	65	6	390	49	6	294
F	182	8	1,456	167	8	1,336
G	38	10	380	45	10	450
H	109	8	872	98	8	784
I	53	6	318	56	6	336
J	139	8	1,112	128	10	1,280
K	126	8	1,008	84	8	672
L	87	10	870	85	10	850
M	114	8	912	111	8	888
N	46	8	368	74	8	592
O	83	6	498	86	6	516
P	124	5	620	121	5	605
Q	83	5	415	77	5	385
R	137	10	1,370	130	10	1,300
Total	1,732		12,650	1,681		12,814

^{a)}A to R used to maintain anonymity of College, ^{b)}Absent students were excluded.

미치며, 이러한 절차들은 하나로 통일되어 있기보다는 매우 다양하다[9,10].

이 연구는 질적 연구방법의 하나인 내용분석(content analysis) 방법을 활용하였다. 양적 연구방법이 가설로부터 시작해 이를 검증하는 방식으로 연구가 이루어지는데 반해, 내용분석 방법은 비구조화된 탐구대상 자료가 함축하고 있는 메시지의 특성을 특정한 분석 준거나 관점에 근거하여 해석함으로써 자료의 특성을 체계적, 객관적으로 규명하는 방법이다[11,12]. 내용분석을 위한 단위(준거)로는 단어, 주제, 인물, 문단, 항목, 개념, 의미 등이 일반적으로 쓰이는데[13], 이 논문에서는 표준화 환자가 기록한 단어와 의미를 준거로 하였다.

2) 연구 절차

연구수행 절차는 자료의 간소화, 자료의 성격 파악 및 코딩, 내용분석, 의미파악의 4단계 과정을 따랐다.

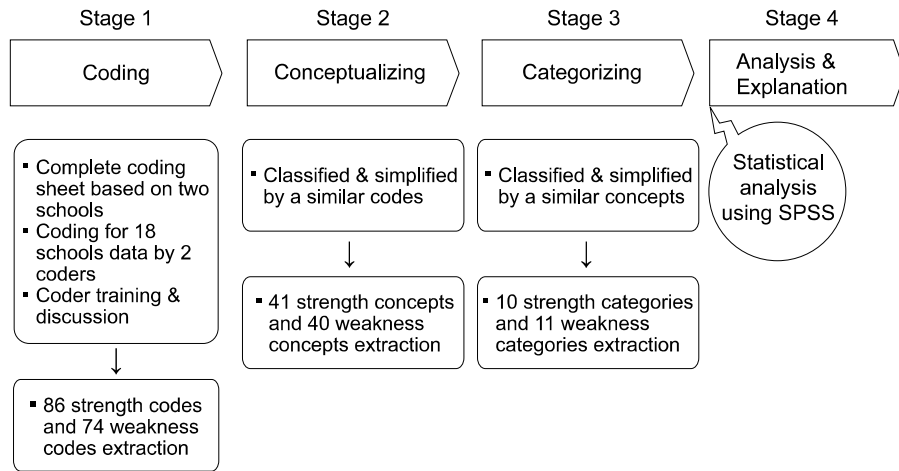
a. 1단계(자료의 간소화): 연도별, 대학별로 입력되어 있는 CPX 채점표 원자료(엑셀파일)는 각 파일의 크기가 매우 방대

했으므로, 이를 단순화하기 위해 '학생의사가 잘한 점'과 '학생 의사가 앞으로 고쳐야 할 점'에 대한 서술형 평가결과와 외의 자료, 즉 '전체만족도', '인사', '병력', '정보나누기', '임상예절', '환자교육' 및 '신체진찰'의 데이터를 파일에서 삭제하였다.

b. 2단계(자료의 성격 파악 및 분석범위 설정): '잘한 점'과 '고쳐야 할 점'에 관한 서술형 평가 내용은 면담과정에서 학생 의사가 표준화 환자에게 보여준 '행동(fact)'과 그 행동으로 인해 환자가 느낀 '감정(feeling)'에 대한 기술로 구분되었다. 학생의사의 행동에 대해 환자가 느낀 감정은 표준화 환자마다 그 감정을 표현하기 위해 사용한 단어의 주관성과 표준화 환자 개인별 편차가 컸기 때문에 분석에서 배제하였고 학생의사가 보여준 '행동'에 관한 내용만 분석하였다.

c. 3단계(코딩): 학생 각 개인에 대한 서술형 평가는 아무런 기술도 되어있지 않은 것부터 한 문장에 환자-의사 관계의 여러 측면을 복합적으로 기술한 것까지 다양했다. 가장 복잡한 내용을 기준으로 했을 때, 학생 1인에 대해 최대 5개까지 최소단위행동(codes)을 추출할 수 있었다. 따라서 학생 1명의

Fig. 1. Analysis Process and Procedure



1개 사례에 대한 표준화 환자의 서술형 평가에 대해 최소 0개에서 최대 5개까지 최소단위행동을 추출하였다. 최소단위행동은 학생의사가 보여준 커뮤니케이션 관련 행동의 가장 세분화된 단위이며 개념 추출을 위한 토대가 되었다. 최소단위행동 추출 방법을 예시하면 다음과 같다.

〈예 1〉 “아무 말씀도 안 하시고 가만히 계실 때 당황했다. 뭐라도 계속 물어보시고 앞으로 어떻게 치료할 건지 왜 이렇게 된 건지 설명해 주실 거라 믿었던만 약을 써보자고 한 이후에 침묵하실 때에는 이해가 잘 안 됐다. 그리고 조금 혀 짧은 발음들이 들려서 신뢰감이 좀 부족한 듯 느껴진다.”
 ⇒ ‘부적절한 침묵’, ‘치료방법에 대한 설명 부족’, ‘비언어적 행동(발음)’의 3개 최소단위행동 추출

〈예 2〉 “일단 가운이 청결하지 않고 면도도 제대로 하지 않은 상태인데 구취까지 좀 심해서 안 그래도 머리가 아픈데 신경이 쓰였다. 좀 물어보시다가 갑자기 침묵하실 때는 뭘 어떻게 해야 하는지 당황스러웠고 계속 MRI, MRA 검사를 해보고 그 결과를 봐야 얘기할 수 있다고 하셨는데 왜 해야 하는지 모르겠다.”
 ⇒ ‘부절절한 용모(구취, 가운 청결)’, ‘부적절한 침묵’, ‘신체 검사하는 이유 설명하기’의 3개 최소단위행동 추출

d. 4단계(내용분석 및 의미파악): 코딩이 완료된 자료는 내용분석 방법에 따라 개념화(conceptualizing), 범주화 또는 유목화(categorizing)의 과정을 통해 결과를 도출하고 설명

및 해석(theorization)을 하였다. 각 단계별로 추출된 최소단위행동, 개념(concepts) 및 범주(categories)에는 고유번호를 지정한 후 자료를 수량화함으로써 체계화하였고 개념화와 범주화 및 유목화가 가능하도록 하였다. 연구의 절차와 분석과정을 요약하면 Fig. 1과 같다.

3) 분석 방법

먼저 연구자가 직접 2개 대학의 서술형 평가 자료를 분석한 후, 최소단위행동에 잘한 점과 개선할 점에 대한 코딩시트(coding sheet)를 작성하였고, 이것을 토대로 2명의 자료분석원(coder)이 18개 대학 자료를 코딩했다. 코딩 과정에서 새롭게 형성된 최소단위행동은 코딩시트에 추가하였고, 유사한 것들은 병합하면서 최종적으로 장점 86개, 단점 74개의 최소단위행동을 추출하였다. 새로운 추가하거나 병합할 최소단위행동에 대한 점점을 위해 자료분석원과 연구자 간에 수차례 논의 과정을 거쳤다.

최종 추출된 커뮤니케이션의 최소단위행동에 대해 유사한 의미를 갖는 것끼리 유목을 구성함으로써 ‘개념’을 추출한 결과, 86개의 ‘잘한 점’ 코드는 41개 개념으로(Table 2), 74개의 ‘고쳐야 할 점’ 코드는 40개 개념(Table 3)으로 수렴되었다. 개념들 간의 속성과 관계를 비교하여 다시 유사 개념끼리 유목화함으로써 최종 범주를 도출하였다. 범주는 연구 결과의 설명 및 해석을 위한 틀로서 학생의사가 보여준 행동의 장·단점 비교분석이 가능해야 했으므로 ‘잘한 점’ 범주와 ‘개선할 점’ 범주를 가능한 유사한 군으로 묶고자 노력하였다. 그 결과 최종

Table 2. Categories and Concepts: Strengths

Categories	Concepts	No. of concepts: total 41
1. Adequate interview atmosphere	Greetings, introduce doctor, appropriate name calling, positive atmosphere	4
2. Active listening	Listening, reaction, facilitation, summarizing	4
3. Emotional support	Express understanding and empathy, reassurance, show interest, respect, acceptance	5
4. Non-verbal behaviors	Voice (tone, accent, pace, pronunciation, volume, etc.), eye contact, expression, face cleanness, speech, touch, silence, writing medical records	8
5. Professional attitude	Professional attitude, favorable attitude, good clinical etiquette	3
6. Questioning	Check reason for visit, questions on symptoms or treatment plans, psychosocial questions, check for patients' understandings, detailed and systematic questions	5
7. Explanation	Detailed explanations, careful and calm explanations, easy explanations, explanations by some cases, answers to the patient's questions, logical and clear explanations, explanations related to physical exams	7
8. Reaching agreement	Decisions on treatment reflected the patients' opinions	1
9. Counseling & education	Patient education, persuades patients, utilize available resources	3
10. Adequate physical examination	Good physical examination skills	1

적으로 '잘한 점' 범주 10개, '개선할 점' 범주 10개를 공통적인 범주로 추출할 수 있었고, '개선할 점' 범주에만 1개 범주가 더 추가되었다(Tables 2, 3). 각 코드별, 개념별, 범주별로 각각 부여한 고유번호에 따라 코딩이 완성된 자료덩이(dataset)는 SPSS WIN version 12.0 (SPSS Inc., Chicago, USA)을 통해 다중응답분석(multiple response analysis)을 시행하였다.

결과

1. 기술통계량

분석대상 사례 수는 2006년 12,650개, 2007년 12,814개였다. 2006년의 경우, 표준화 환자의 서술형 피드백이 한 가지 이상 기록된 사례는 9,104개(72.0%)였고, 한 마디도 기록되지 않은 사례는 3,546개(28.0%)였다. 2007년에는 '잘한 점' 또는 '고쳐야 할 점'이 한 가지 이상 기록되었던 사례가 8,470개(66.1%)였고, 아무런 기록도 없었던 사례가 4,344개(33.9%)였다. '잘한 점'과 '고쳐야 할 점' 두 가지를 한 가지 이상 동시에 기록한 사례는 2006년에 2,495개(19.8%), 2007년에 2,336개(18.2%)였다. '잘한 점'과 '고쳐야 할 점' 중 한 가지만 기록

한 것 중에는 '잘한 점'만을 기록한 사례가 더 많았다(2006년 잘한 점만 기록된 사례는 28.8%, 고쳐야 할 점만 기록된 사례는 23.4%; 2007년 잘한 점만 기록된 사례는 32.2%, 고쳐야 할 점만 지적받은 사례는 15.7%였음). 2006년과 2007년 모두 표준화 환자는 환자-의사 관계 측면에서 학생의사가 고쳐야 할 점보다 잘한 점을 더 많이 기록하였다(Table 4).

2. 환자-의사 커뮤니케이션 범주별 분석 결과

1) 학생의사가 '잘한 점'과 '고쳐야 할 점' 범주와 하위 개념들

내용분석 방법에 따른 체계적 분석과정을 거쳐 표준화 환자의 서술형 평가로부터 추출한 환자-의사 관계의 내용 범주는 '잘한 점' 10개, '고쳐야 할 점' 11개였다. 잘한 점과 고쳐야 할 점의 공통적 범주는 10개로서 '면담분위기', '경청하기', '정서적 지지', '비언어적 행동', '태도', '질문하기', '설명하기', '합의도달', '환자교육 및 상담', '신체검진능력'이었다. 고쳐야 할 점 범주에만 한 가지 추가된 것은 '체계적, 논리적 면담진행'이었다(Tables 2, 3). 범주 추출을 위한 하위 수준의 준거는 개념이었는데, 최종 추출된 개념 수는 '잘한 점'이 40개, 고쳐야 할 점이 41개였다.

10개의 '잘한 점' 범주에 포함된 하위 개념 40개를 구체적으

Table 3. Categories and Concepts: Weaknesses

Categories	Concepts	No. of concepts: total 40
1. Adequate interview atmosphere	Did not greet, inappropriate name calling, inappropriate mood	3
2. Active listening	Does not listen well (cut off patients speaking, failure to remember, did not listen carefully, did not listen to the end), does not summarize	2
3. Emotional support	Failure to agree and empathy, did not reassurance or console, indifference, no caring, not acceptable	5
4. Non-verbal behaviors	Voice (tone, accent, pace, pronunciation, volume, etc.), eye contact, expression, face cleanness, speech, distance, inappropriate pose and action, silence, inappropriate medical records	9
5. Professional attitude	Lack of professional attitude, Businesslike authoritative attitude, lack of authenticity, lack of clinical manner	4
6. Questioning	Does not check reason for visit, lack of questions related to symptoms, lack of psychosocial questions, does not check for patients' understandings, unstructured questions, closed questions	6
7. Explanation	Inadequate explanations, inappropriate explanations, failure to make easy explanations, failure to diagnose notification, does not speaking of doctors' opinion	5
8. Reaching agreement	Unilateral treatment decisions, hasty conclusions	2
9. Counseling & education	Failure to take advantage of resources	1
10. Adequate physical examination	Unskilled physical examinations	1
11. Structured & organized interview	Unstructured medical interview, hasty medical interview	2

Table 4. Descriptive Statistics

Categories	2006	2007
	Case, No. (%)	Case, No. (%)
Describe strengths only	3,647 (28.8)	4,125 (32.2)
Describe weaknesses only	2,962 (23.4)	2,009 (15.7)
Both strengths and weaknesses	2,495 (19.8)	2,336 (18.2)
Non-description	3,546 (28.0)	4,344 (33.9)
Total	12,650 (100.0)	12,814 (100.0)

로 살펴보면, '면담분위기' 4개(인사하기, 자기소개, 적절한 호칭, 긍정적 분위기 형성), '경청하기' 4개(듣기, 반응해주기, 촉진하기, 요약하기), '정서적 지지' 5개(이해와 공감 표현, 안심과 위로, 관심표현, 배려, 수용), '비언어적 행동' 8개(목소리·역양·속도·발음·크기, 눈맞춤, 표정, 용모·청결, 말투 신체접촉, 침묵, 의무기록 작성), '태도' 3개(의사다운 태도, 호의적인 태도, 임상예절), '질문하기' 5개(내원이유 확인, 증상 또는 치료계획에 대한 질문, 사회심리적 질문, 환자의 이해정도를 확인하는 질문, 체계적이고 꼼꼼한 질문), '설명하기' 7개

(상세한 설명, 신중하고 차분한 설명, 쉬운 설명, 예를 들어 설명, 환자질문에 대한 답변, 논리적이고 명확한 설명, 신체진찰 관련 설명), '합의도달' 1개(환자 의견을 반영한 치료방법 결정), '환자교육 및 상담' 3개(환자교육, 환자설득, 주위자원 활용), '신체검진능력' 1개(신체진찰을 잘함)였다.

11개의 '고쳐야 할 점' 범주에 포함된 하위 개념 41개는 '면담분위기' 3개(인사하지 않음, 부적절한 호칭, 부적절한 분위기), '경청하기' 2개(잘 듣지 않음-예: 환자가 말하는 중간에 끼어들기·기억 못함·귀담아 듣지 않음·끝까지 듣지 않음,

요약하지 않음), '정서적 지지' 5개(이해와 공감 못함, 안심과 위로 못함, 무관심, 배려 못함, 수용 못함), '비언어적 행동' 9개(목소리-톤·억양·속도·발음·크기, 눈맞춤 못함, 좋지 못한 표정, 용모불결, 부적절한 말투, 거리, 부적절한 자세나 동작, 부적절한 침묵, 의무기록 작성 안함), '태도' 4개(전문가적 태도 부족, 사무적·권위적 태도, 진정성 결여, 임상예절 부족), '질문하기' 6개(내원 이유 확인 안함, 증상관련 질문 부족, 사회심리적 질문 부족, 환자의 이해정도를 확인하는 질문 부족, 애매한 또는 한꺼번에 많은 부적절한 질문, 폐쇄형 질문), '설명하기' 5개(불충분한 설명, 부적절한 설명, 쉬운 설명 못함, 진단통보 미숙, 의사생각을 환자에게 말하지 않음), '합의도달' 2개(일방적 치료결정, 성급한 결론), '환자교육 및 상담' 1개(주위자원 활용 못함), '신체검진능력' 1개(신체진찰 미숙), '논리적 체계적 면담진행' 2개(체계성 부족, 성급한 면담 진행)였다.

2) '잘한 점'과 '고쳐야 할 점'의 범주별 다중응답분석 결과

코드화, 개념형성, 범주화 과정을 거쳐 최종 분석 틀로 확정된 '잘한 점' 범주 10개와 '고쳐야 할 점' 범주 11개를 기준으로 학생의사의 강점과 취약점을 분석한 결과는 Table 5와 같다.

2006년의 경우 학생의사가 잘한 점의 총 개수는 11,147개였고, 고쳐야 할 점의 총 개수는 7,768개였다. 학생들은 표준

화 환자로부터 '고쳐야 할 점'보다는 '잘한 점'을 약 1.5배 정도 더 많이 평가받았다. 2007년에도 고쳐야 할 점보다 잘한 점에 대한 서술형 평가가 약 2배 정도 많았는데, 잘한 점의 총 개수는 11,386개, 고쳐야 할 점의 총 개수는 5,875개였다.

환자-의사 관계 측면에서 표준화 환자로부터 학생의사가 '잘했다'는 평가를 받은 범주를 상위 순서대로 보면, 2006년에는 정서적 지지(3,011개, 27.0%), 설명하기(2,325개, 20.9%), 태도(1,993개, 17.9%), 면담분위기(1,423개, 12.8%), 비언어적 행동(718개, 6.4%), 경청(676개, 6.1%), 환자교육 및 상담(663개, 5.9%), 질문하기(193개, 1.8%), 신체검진능력(102개, 0.9%), 합의도달(35개, 0.3%) 순으로서, 1에서 3순위(정서적 지지, 설명하기, 태도)까지가 잘한 점 전체의 65.8%를 차지했다. '고쳐야 할 점'의 범주별 상위 평가 순위는 비언어적 행동(2,674개, 34.4%), 설명하기(1,888개, 24.3%), 태도(1,552개, 20.0%), 정서적 지지(448개, 5.8%), 신체검진능력(334개, 4.3%), 질문하기(283개, 3.6%), 합의도달(246개, 3.2%), 경청하기(151개, 1.9%), 체계적 면담진행(96개, 1.2%), 면담분위기(91개, 1.1%), 환자교육 및 상담(14개, 0.2%) 순으로서, 1순위에서 3순위(비언어적 행동, 설명하기, 태도)까지가 단점 전체의 78.7%나 되었다(Table 5).

2007년에도 학생의사가 잘한 점과 고쳐야 할 점의 평가 순위는 크게 변화되지 않았다. 즉, '잘한 점'으로 높은 평가를 받

Table 5. Strengths and Weaknesses by Categories in 2006 and 2007

Categories	2006		2007	
	Strengths	Weaknesses	Strengths	Weaknesses
1. Adequate interview atmosphere	1,423 (12.8)	82 (1.1)	1,387 (12.2)	85 (1.4)
2. Active listening	676 (6.1)	151 (1.9)	821 (7.2)	172 (2.9)
3. Emotional support	3,011 (27.0)	448 (5.8)	2,494 (21.9)	774 (13.2)
4. Non-verbal behaviors	718 (6.4)	2,674 (34.4)	727 (6.4)	1,870 (31.8)
5. Professional attitude	1,993 (17.9)	1,552 (20.0)	2,211 (19.4)	1,178 (20.1)
6. Questioning	201 (1.8)	283 (3.6)	193 (1.7)	108 (1.8)
7. Explanation	2,325 (20.9)	1,888 (24.3)	2,811 (24.7)	1,201 (20.4)
8. Reaching agreement	35 (0.3)	246 (3.2)	86 (0.8)	250 (4.3)
9. Counseling & education	663 (5.9)	14 (0.2)	535 (4.7)	33 (0.6)
10. Adequate physical examination	102 (0.9)	334 (4.3)	121 (1.1)	140 (2.4)
11. Structured & organized interview	NA	96 (1.2)	NA	64 (1.1)
Total	11,147 (100.0)	7,768 (100.0)	11,386 (100.0)	5,875 (100.0)

Values are presented as number (%).

NA: Not available.

은 상위 3순위로서 설명하기(2,811개, 24.7%), 정서적 지지(2,294개, 21.9%), 태도(2,211개, 19.4%)는 1, 2순위의 변동만 있을 뿐 잘한 점 전체의 66.0%를 차지함으로써 전년도 65.8%와 대동소이한 수준이었다. '고쳐야 할 점'에 있어서도 연도에 따른 뚜렷한 변화가 발견되지 않았는데 2007년도에 가장 많은 지적을 받은 3개 범주, 즉 비언어적 행동(1,870개, 31.8%), 설명하기(1,201개, 20.4%), 태도(1,178개, 20.0%)는 전년도와 마찬가지로 순위변동 없이 1순위에서 3순위를 차지했다. 3순위까지의 빈도는 '고쳐야 할 점' 전체 빈도의 82.4%나 되어서 전년도의 78.7%보다 상승한 결과를 보여주었다 (Table 5).

이상의 결과를 종합해 보면, 표준화 환자의 평가에 근거해 볼 때 CPX 상황에서 학생들이 수행하는 환자-의사 관계 행동 중 '정서적 지지'와 '면담분위기' 같은 커뮤니케이션 행동은 '고쳐야 할 점'으로서 보다는 '잘하는 점'으로 평가된다. 그러나 '비언어적 행동'은 '잘한 점'으로서가 아니라 '고쳐야 할 점'으로 훨씬 더 많은 지적을 받고 있다. '설명하기'와 '태도'는 2006년과 2007년 공통적으로 학생의사가 '잘한 점'은 물론 '고쳐야 할 점'으로도 각각 20% 내외에 달하는 많은 평가를 받고 있어서 환자-의사 관계에서 평가에 중요한 영향을 미치는 것으로 확인되었다. 그 밖에, 전체 빈도 중 차지하는 비율은 낮지만 '고쳐야 할 점'이기보다 '잘한 점'으로 더 많은 평가를 받

은 항목은 '경청하기'와 '환자교육 및 상담'이었고, 반대로 '잘한 점'이기보다 '고쳐야 할 점'으로 더 많이 지적된 항목은 '질문하기', '합의도달', '신체검진' 항목이었다.

3. CPX 사례별 분석 결과

2006년과 2007년의 각 대학별 CPX 시험 사례 수는 최소 5개에서 최대 10개까지 대학별 편차가 있었기 때문에 여기서는 모든 대학에서 공통적으로 시험을 치렀던 사례 중 2006년의 사례 1, 사례 2, 사례 3 결과만을 비교 분석하였다.

환자-의사 관계 측면에서 학생의사가 보여 준 '잘한 점'과 '고쳐야 할 점'의 빈도는 사례별로 분명한 차이가 있었다. 2006년의 경우 '잘한 점'이 가장 많이 기술된 것은 사례 1 (1,725개), 사례 3 (1,313개), 사례 2 (797개) 순이었고, '고쳐야 할 점'으로 많은 지적을 받은 것은 그 반대 순서인 사례 2 (1,382개), 사례 3 (782개), 사례 1 (649개) 순이었다(Table 6). 사례 1은 사례 3보다 '잘한 점'에서 2배 이상 더 많은 긍정적 평가를 받은 반면, 사례 2는 '고쳐야 할 점'에서 사례 1보다 2배가량 더 많은 부정적 평가를 받았다.

'신체진찰'과 '임상의사로서의 예절'이 강조되어 있는 2006년의 사례 1은 '태도를 '잘한 점(24.0%)'으로 평가하거나 '고쳐야 할 점(23.9%)'으로 지적한 빈도가 사례 2보다 훨씬 높았다(사례 2의 잘한 점 14.9%, 고쳐야 할 점 15.9%). 반면에 '정

Table 6. Strengths and Weaknesses by Case 1, Case 2, and Case 3 in 2006

Categories	Strengths			Weaknesses		
	Case1	Case 2	Case 3	Case1	Case 2	Case 3
1. Adequate interview atmosphere	213 (12.3)	77 (9.7)	226 (17.2)	3 (0.5)	20 (1.4)	6 (0.8)
2. Active listening	181 (10.5)	66 (8.3)	40 (3.0)	5 (0.8)	35 (2.5)	7 (0.9)
3. Emotional support	243 (14.1)	225 (28.2)	284 (21.6)	19 (2.9)	89 (6.4)	23 (2.9)
4. Non-verbal behaviors	219 (12.7)	38 (4.8)	45 (3.4)	280 (43.1)	394 (28.5)	186 (23.8)
5. Professional attitude	414 (24.0)	119 (14.9)	332 (25.3)	155 (23.9)	220 (15.9)	214 (27.4)
6. Questioning	13 (0.8)	20 (2.5)	21 (1.6)	36 (5.5)	51 (3.7)	34 (4.3)
7. Explanation	238 (13.8)	225 (28.2)	328 (25.0)	65 (10.0)	456 (33.0)	138 (17.6)
8. Reaching agreement	3 (0.2)	4 (0.5)	—	16 (2.5)	80 (5.8)	54 (6.9)
9. Counseling & education	157 (9.1)	22 (2.8)	27 (2.1)	—	12 (0.9)	—
10. Adequate physical examination	44 (2.6)	1 (0.1)	10 (0.8)	64 (9.9)	—	100 (12.8)
11. Structured & organized interview	NA	NA	NA	6 (0.9)	25 (1.8)	20 (2.6)
Total	1,725 (100.0)	797 (100.0)	1,313 (100.0)	649 (100.0)	1,382 (100.0)	782 (100.0)

Values are presented as number (%).
NA: Not available.

신과적 평가 및 정신상태 평가를 포함하고 있는 사례 2는 '정서적 지지' 면에서 사례 1보다 더 많은 서술형 평가를 받았는데, '잘한 점'으로 225개(28.2%), '개선할 점'으로 89개(6.4%)를 평가받았다. '설명하기'도 사례 2가 사례 1보다 '잘한 점' 또는 '고쳐야 할 점'에서 더 많은 평가를 받았다. 즉, '정서적 지지'와 '설명하기'는 사례 1보다 사례 2에서 더 빈번하게 평가되었고, '태도'는 사례 2보다 사례 1에서 더 높은 빈도로 평가되었다.

비언어적 행동은 '잘한 점'으로서 보다는 '고쳐야 할 점'으로 더 자주 평가되었는데, 특히 사례 1은 '고쳐야 할 점' 전체 빈도 중 43.1%가 비언어적 행동에 대한 지적이었다. 그 밖에 '신체진찰'과 '경청하기'도 사례 2보다 사례 1에서 장점 또는 개선할 점으로 더 빈번하게 평가되었다.

이상의 결과를 종합해 보면, 환자-의사 관계 측면에서 학생들이 잘하는 것과 취약한 점은 CPX 사례에 따라 달라질 수 있다. 즉, 사례 1에서는 '태도'가 두드러진 강점으로, '비언어적 행동'이 가장 취약한 점으로 나타나지만, 사례 2에서는 '정서적 지지'가 상대적인 강점을 가지는 반면 '설명하기'는 더 큰 약점으로 부각되고, 사례 3에서는 '태도'가 다른 사례에 비해 더 취약한 것으로 보인다(Table 6).

고찰

이 연구는 CPX 평가체제에서 표준화 환자가 자유기술형 피드백 내용을 체계적으로 분석하여 환자-의사 관계 측면에서 학생의사의 강점과 취약점을 파악하고, 정량적 평점표에 비하여 서술형 평가가 환자-의사 관계 영역의 수행능력에 대해 추가적으로 제공하는 정보가 무엇인지를 확인함으로써 피드백 장치로서 서술형 평가의 교육 평가적 의미와 기능을 점검하고자 하였다.

질적 연구방법의 하나인 내용분석 방법에 따라 서울·경기 CPX 컨소시엄 소속 18개 대학의 2006년과 2007년 자료를 분석한 결과, 환자-의사 관계 측면의 내용으로서 '잘한 점' 범주 10개(41개 개념 포함), '고쳐야 할 점' 범주 11개(40개 개념 포함)를 추출하였다(Tables 1, 2).

이들 범주와 개념을 CPX 평가표의 구조화된 채점 항목 7

개와 각 항목별 세부기술 내용과 비교해 본 결과, 자유기술형 평가내용은 구조화된 평가표가 제시하는 모든 항목을 포괄하고 있었고, 추가적으로 구조화된 평가지가 포괄하지 못하는 세밀하고 심층적인 정보까지 담고 있음이 발견되었다. 구조화된 평가표에는 없었지만 자유기술형 피드백 내용 속에서 추출되었던 커뮤니케이션 항목은 '적절함(부절적함) 호칭', '안심시켜 줌(못줌)', '말투', '신체접촉(touch)', '침묵', '호의적인(술직, 성실, 진지 등) 태도', '내원이유 확인', '환자의 이해 정도를 확인하는 질문', '부적절한 질문(예: 애매한 질문, 한꺼번에 많은 질문)', '신중한, 차분한 설명', '환자질문에 대한 설명', '진단통보 미숙', '부적절한 설명(한꺼번에 많은 설명, 엉뚱한 설명, 똑같은 설명 반복, 말을 돌리면서 정확히 설명하지 않음)', '성급한 결론', '체계적 부족' 등이었다. 이러한 항목들은 CPX 상황에서 환자-의사 관계 측면의 구조화된 채점표를 보완할 수 있는 구체적인 피드백 정보로 활용될 수 있다.

현재 CPX 평가표의 서술형 평가 결과는 구조화된 평가 결과와 함께 학생들 개개인에게 통보하도록 하고 있다. 구조화된 평가 결과는 해당 대학의 환자면담 영역별, 사례별 최고점과 최하점, 그리고 학생개인이 획득한 점수를 보여주어 사례별 합격/불합격 여부와 전체 그룹 속에서 학생 개인의 상대적 위치를 보여준다. 환자-의사 관계의 양적 평가는 학생의 성취도에 대한 진단적 혹은 총괄적 평가 기능을 제공하지만 형성적 평가로서의 기능을 수행하는데 한계가 있다. 반면 각 시험 사례별로 제시되는 '잘한 점'과 '개선할 점'에 대한 서술형 평가는 시험상황에서 학생의사가 사례별로 수행했던 각각의 행동에 대해 표준화 환자가 느꼈던 생각과 감정을 구체적으로 아주 생생하게 전달하기 때문에 단일 점수로 표기되는 구조화된 평가보다 개별적이고 구체적인 정보를 제공한다. 좋은 피드백은 양질의 정보를 제공함으로써 학습에 있어서 자기반성을 촉진시키고, 현재의 수행과 바람직한 수행 간의 격차를 줄일 수 있게 해 주므로[14,15,16], 서술형 평가는 양적 평가와는 다른 측면에서 다양하고 세밀한 정보를 제공하여 학생들이 자신의 수행을 점검해 볼 수 있는 유용한 피드백 기능을 수행한다고 볼 수 있다.

한편, 환자중심적 의료면담을 수행하는데 있어 중요한 커뮤니케이션 기술들이 표준화 환자의 서술형 평가에서도 중요하게 인식되고 있음이 확인되었다. 즉, CPX 상황에서 표준화

환자로부터 '잘한 점'과 '고쳐야 할 점'의 상위 3순위 범주에 속했던 '설명하기', '정서적 지지', '태도', '비언어적 행동'의 항목들과 잘한 점으로 더 많은 평가를 받았던 '경청하기'와 '환자교육 및 상담', 그리고 고쳐야 할 점으로 더 많이 지적되었던 '질문하기', '합의도달'은 환자중심적 커뮤니케이션의 중요 항목들로서, Cole & Bird [17]가 주장했던 의학면담의 세 가지 기본적 기능, 즉 라포 형성하기, 좋은 면담자료 수집하기, 순응도 높이기를 달성하기 위해 필수적으로 갖추어야 할 기술들이다.

이상의 결과를 통해 다음과 같은 결론을 내릴 수 있다.

첫째, 학생들에게 상대적으로 취약한 커뮤니케이션 기술은 있으며, 이러한 기술들에 대한 교육이 의료 커뮤니케이션 수업에서 강화될 필요가 있다. 이 연구에서 '개선할 점'으로 가장 많은 지적을 받은 '비언어적 행동' 교육이 특히 강화되어야 한다. 말은 포장된 생각을 전달할 수 있지만 비언어적 행동은 솔직한 생각을 표현하기 때문에 언어보다 더 많은 것을 전달한다[18]. 의사가 환자와 커뮤니케이션을 잘하기 위해 경청하고 공감적 커뮤니케이션을 실천해도, 비언어적 커뮤니케이션이 적절하지 않으면 그의 의도가 온전히 전달되지 못할 수 있다. 이러한 기술들은 학생 자신이 스스로 인식하기 어렵고 단기간에 변화를 이끌어 내기도 쉽지 않기 때문에 반복된 훈련 기회를 제공함으로써 지속적으로 개발시켜야 한다.

둘째, 환자-의사 관계 측면에서 학생들이 '잘하는 것'과 '취약한 점'은 CPX 사례에 따라 달라질 수 있으므로, 교육장면에서 다양한 사례 개발과 사례 체험 교육 기회가 확대될 필요가 있다. 신체검진이 어려운 사례, 두통이나 감기 같은 흔한 증상에 기초한 진단 결정 사례, 설명이 어려운 사례, 우울, 갱년기 장애, 스트레스 등 심리사회적 정보수집이 필요한 사례 등 다양한 사례를 통해 학생들의 커뮤니케이션 교육의 효과를 높여야 한다.

셋째, 이 연구에서 추출된 장·단점 범주는 7개 항목으로 구성된 구조화된 환자-의사관계 평점표가 담지 못하는 보다 세부적인 정보를 보여주었다. 따라서 서술형 평가 자료가 구조화된 양적 평가표의 보완기능을 수행하는 형성평가도구로써 적극 활용되도록 해야 한다. 그러기 위해서는 표준화 환자가 서술형 평가의 의미를 명확히 인식하고, 평가를 함에 있어서 학생의사의 임상수행능력 개선에 도움이 되도록 피드백

할 수 있는 방법과 기술을 갖추도록 교육해야 한다. 표준화 환자는 평가해야 할 행동의 기준(criteria)과 표준(standards)에 대한 깊이 있는 지식을 가져야 하고[15,16], 학생들을 돕고자 하는 마음과 자신이 기술한 내용이 타당한지에 대한 관심, 자기만 독특한 피드백 패턴(예: 잘한 점을 덜 평가하고, 고쳐야 할 점을 더 많이 평가하는 경향 등) 등에 대해 인식해야 한다.

이 연구의 제한점은 질적 자료로서 분석대상 자료가 갖는 한계를 갖고 있다는 점이다. 즉, 이 연구는 연구 대상 자료가 표준화 환자의 주관적 판단에 따라 특정 항목이 더 빈번하게 기록되거나, 반대로 덜 기술된 문제를 극복하지 못하는 한계가 있다. 따라서 향후 연구에서는 표준화 환자 개인별 선호 항목을 고려한 분석이 추가될 필요가 있다. 또한 추가적인 연구로서 표준화 환자의 피드백을 통해 학생들이 자신의 커뮤니케이션 행동에 대해 어떤 새로운 생각과 변화를 가지게 되었는지에 대한 연구도 진행되어야 할 것이다.

결론적으로, CPX 평가체제에서 환자-의사 관계평가는 서술형 평가가 갖는 보완적 기능을 최대한 활용함으로써 학생들의 임상수행 결과에 대한 보다 체계적인 피드백을 제공할 수 있을 것이다. 서술형 피드백의 목적은 환자-의사 관계에 필요한 지식과 기술을 발달시키는 것뿐만 아니라, 학생들이 자기 주도적 학습을 할 수 있는 힘을 길러주기 위한 방법으로 사용되어야 한다.

Acknowledgements: We would like to thank for giving data to the Seoul-Gyeonggi CPX consortium. Also, we would like to express our appreciation to the TA, Hei Lim Kim.

Funding: None.

Conflicts of interest: None.

REFERENCES

1. Norman GR, Neufeld VR, Walsh A, Woodward CA, McConvey GA. Measuring physicians' performances by using simulated patients. J Med Educ 1985; 60: 925-934.

2. Hodges B, Turnbull J, Cohen R, Bienenstock A, Norman G. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ* 1996; 30: 38-43.
3. Wallace J, Rao R, Haslam R. Simulated patients and objective structured clinical examinations: review of their use in medical education. *Adv Psychiatr Treat* 2002; 8: 342-348.
4. Ladyshevsky R. Simulated patients and assessment. *Med Teach* 1999; 21: 266-269.
5. Klamen DL, Yudkowsky R. Using standardized patients for formative feedback in an introduction to psychotherapy course. *Acad Psychiatry* 2002; 26: 168-172.
6. Kim S, Park S, Hur Y, Lee SJ. The appropriateness of using standardized patients' (SPs) assessment scores in clinical performance examination (CPX). *Korean J Med Educ* 2005; 17: 163-172.
7. Kwon I, Kim N, Lee SN, Eo E, Park H, Lee DH, Park MH, Oh J, Han JJ, Huh J, Ryu KH. 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.
8. Han JJ, Park H, Kwon I, Ryu KH, Eo E, Kim N, Jung J, Kim KH, Lee SN. The comparison of clinical performance examination scores according to the different testing time: six medical schools in seoul. Gyeonggi CPX consortium 2005. *Korean J Med Educ* 2007; 19: 31-38.
9. Creswell JW. *Qualitative inquiry and research design: choosing among five traditions*. London, UK: Sage Publications; 1997.
10. Creswell JW. *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, USA: Sage Publications; 2003.
11. Krippendorff K. *Content analysis: an introduction to its methodology*. 2nd ed. Thousand Oaks, USA: Sage Publications; 2004.
12. Kim YC. *Methods and processes of qualitative research for inquiring educational phenomenon in school*. *Korean J Educ Res* 1997; 35: 135-170.
13. Berelson B. *Content analysis in communication research*. New York, USA: Hafner; 1952.
14. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud High Educ* 2006; 31: 199-218.
15. Sadler DR. Formative assessment: revisiting the territory. *Assess Educ* 1998; 5: 77-84.
16. Bose J, Rengel Z. A model formative assessment strategy to promote student-centered self-regulated learning in higher education. *US-China Educ Rev* 2009; 6: 29-35.
17. Cole SA, Bird J. *The medical interview: the three-function approach*. St Louis, USA: CV Mosby; 2000.
18. Platt FW, Gordon GH. *Field guide to the difficult patient interview*. 2nd ed. Philadelphia, USA: Lippincott Williams & Wilkins; 2004.