

채점자의 입실이 OSCE 응시자 성적과 수행에 미치는 영향

영남대학교 의과대학 의학교육학교실

방 재 범

Effects of Rater's Presence in OSCE Station on Examinee's Scores and Performance

Jae Beum Bang

Department of Medical Education, Yeungnam University, College of Medicine, Daegu, Korea

Purpose: It is uncertain whether the rater's presence during administration of the objective structured clinical examination (OSCE) has any impact on examinee performance. To clarify this issue, the effects of a rater's presence during testing on OSCE score and examinee responses were analyzed.

Methods: The OSCE comprised 24 stations of 12 duplicated cases. Twenty-eight of 83 fourth-year medical students were placed in four rater-absent stations (pediatrics, 2 stations; internal medicine, 2 stations). The scores of the rater-absent group were assessed after review of recorded videotapes. We obtained student responses on the rater's presence by means of a questionnaire. Statistical analysis (t-test) was performed using SPSS 14.0.

Results: The scores of the rater-absent group compared with the rater-present group were 78.57 ± 16.27 vs. 70.55 ± 19.57 ($p=0.066$) for pediatrics, 66.07 ± 14.74 vs. 64.36 ± 14.88 ($p=0.621$) for internal medicine, and 67.75 ± 5.43 vs. 69.13 ± 4.85 ($p=0.099$) for total score, respectively. Subjects in the rater-absent group felt that they were in a real-life clinical situation ($p=0.013$) and concentrated better ($p=0.006$), and more students requested to take the OSCE without a rater in the station ($p=0.000$) than those in rater-present group.

Conclusion: Although there was no significant difference in OSCE scores between the rater-absent and rater-present examinee groups, subjects at rater-absent stations felt more of a real clinical situation and could concentrate more effectively than those at rater-present stations. To this end, videotaped OSCE assessments should be considered to increase the validity of the examination.

Key Words: Clinical competence, Educational measurement

서 론

Received: April 23, 2008 Accepted: August 19, 2008

Corresponding Author: Jae Beum Bang
medu@med.yu.ac.kr

* This research was revised, which had been presented at the 19th Symposium of the Korean Society of Medical Education.

진료수행능력에 대한 평가는 장차 의사가 될 학생들이 의사로서 실제 역할을 수행할 수 있는지를 평가하는 것으로서 1975년 의학교육계에 처음 도입되었다 (Harden & Gleeson, 1979). 이러한 평가방식

은 의사로서의 역량을 갖추게 하기 위한 교육과 그러한 역량을 실제 발휘할 능력이 있는지를 평가한다는 점에서 이전의 도제식 교육이나 지필중심의 평가와 같은 전통적 방식에 큰 변화를 야기했다. 최근 우리나라 의학교육제도 학생들의 지식 중심의 교육과 그에 대한 평가에서 벗어나 술기와 태도에 관한 역량을 함양시키는 데에 많은 주의를 기울이고 있다. 특히 2009년부터 의사국가시험에 임상수행능력에 대한 평가가 적용되어질 예정이기에 이러한 평가 방법과 그에 따른 객관성 확보가 매우 중요한 이슈로 떠오르고 있다.

국내에서 이러한 진료수행능력을 평가하는 것과 관련하여 객관구조화진료시험 (objective structured clinical examination, OSCE)에 대한 채점과 그에 대한 신뢰도 검증을 위한 연구가 이루어졌다. OSCE 문항 및 관련 정보 유출을 막기 위해 그리고 채점 교수 동원의 어려움을 해결하기 위해 시험장을 복제하여 많은 학생들을 한 번에 시험을 치르게 한다. 이와 관련하여 Park *et al.* (1999)은 시험장 복제에 따른 신뢰도에 있어서 시험장에 따른 큰 점수 차이는 나지 않는 것으로 확인했다. 채점 교수가 아닌 진료 대상자에 의한 평가의 적합성을 근거로 표준화환자 (standardized patient, SP) 채점의 신뢰도 여부 검증과 관련하여 Park *et al.* (2003)은 교수와 표준화환자 사이의 채점평가표에 대한 일치도 검증에 대해 71~82%로 보고하면서 점검표의 내용과 척도 구간에 대한 명확한 정의와 충분한 채점 훈련이 이루어져야 할 것을 지적했다. Shin *et al.* (2005)은 충분한 훈련 이후의 교수채점자의 채점을 정답으로 간주하고 SP 채점을 검증한 결과 86.9점이라는 높은 정확도를 확인하면서도 이러한 정확도를 떨어뜨리는 요인을 채점표와 채점기준표의 체계가 복잡하거나 복수질문을 하거나 제시된 판단의 준거가 복잡하다는 점을 지적하고 있다.

신뢰도와 관련하여 최근 해외 연구는 특히 점검표 채점 (checklist rating)과 그에 대한 지침에 있어서의 문항개발자와 채점자, 채점자간의 평가지침에 대한 해석과 학생 수행에 대한 평가 입장의 근본적인 차이를 감안하여 글로벌 채점 (global rating)이나

OSCE 중간에 필기시험이나 실제 실습 과정 동안의 평가가 의사 자격을 부여하는 평가에 보완적으로 이루어져야 한다는 것을 강조하고 있다 (Regehr *et al.*, 1998). 최근 진료수행평가의 신뢰도 향상을 위해 채점평가표 (checklist)에 의해서만 평가가 이루어지 기보다는 글로벌 채점 (global rating)에 의한 보완이 강조되고 있다. 실제로 Regehr *et al.* (1998)은 전반적으로 글로벌 채점이 점검표 채점보다 더 신뢰도가 높은 것으로 보고하고 있고 Newble (2004)는 실제적인 임상수기영역은 채점평가표를 통한 평가가 적합하고, 환자면담이나 환자교육과 같은 영역에 대해서는 글로벌 채점방식이 적합하다고 보고하고 있다. 이러한 신뢰도 이외에 타당도와 관련해서 Newble (2004)은 임상수행능력을 평가하는 데에 있어서의 구성요소와 이러한 평가 과정에 대한 청사진과 전반적인 과정 진행도가 필요하다는 것을 강조했다.

수행평가의 경우 실제 수행이 요구되는 상황에 대한 행동적 대처 및 상황과 관련된 문제해결에 초점을 둔다는 점에서 전통적인 지식중심의 평가 방식과 차이가 있다. 일반적으로 지식을 획득했는지 혹은 이해하고 있는지에 대한 평가는 지필고사에 의한 방법을 통해 확인하며 이에 대해서는 평가에 있어서 평가 도구의 신뢰도에 초점을 두지만 수행평가의 경우 평가의 목적에 관한 타당도에 더 많은 강조를 둔다 (Kim, 2000). 이를 의학교육과 관련해서 보면 임상수행평가는 실제적인 임상상황에 가까운 상황을 제시하고 그에 적합한 내용을 중심으로 평가해야 한다는 것을 의미한다. 진료수행평가와 관련하여 타당성에 관한 연구는 최근 평가 그 자체에 대한 논의뿐만 아니라 평가가 이루어지는 상황과 응시자가 실제 의료활동을 하는 상황의 맥락 (context)에 대한 이해와 반영이 강조되고 있다. Hodges (2003a, 2003b)에 의하면 진료수행평가에 있어서 진료 상황에 대한 실제감 (real-life)과 시험의 관련은 내적 타당도 (content validity) 뿐만 아니라 맥락적 충실성 (contextual fidelity)에 있어서 매우 중요하다고 본다. 이러한 점에서 진료수행능력 평가에 대한 시험장 환경은 매우 중요하다고 볼 수 있다. 시험장 환경, 특히 환자 진료 활동에 대한 평가는 응시 학생들이 장차 의사로서 활동하

는 진료상황이어야 하며, 그러한 진료상황적 조건에 근거하여 평가가 이루어져야 한다.

하지만 일반적으로 국내에서 이루어지는 OSCE에서는 평가자가 시험장 즉 스테이션에 입실하여 평가해 온 것이 일반적이다. 하지만 이는 실제 진료 상황과 다른 시험장 환경을 제시하는 것이며 바로 이러한 점이 응시자들에게 또 다른 영향을 줄 수 있다. Kirby & Curry (1982)가 지적한 바대로 이러한 수행평가는 다른 일반적인 지필평가 방법보다 응시자들에게 더 많은 긴장감을 준다는 점을 고려한다면 채점자가 입실한 상태는 학생들에게 긴장감을 더해 줄 수 있다. 실제 임상실습 교육동안 직접적인 환자 예진 경험이 전반적으로 부족하며 진료수행능력 평가에 충분히 익숙하지 않은 우리나라 의학교육 실정을 고려한다면 이러한 시험장 환경 조건은 학생들의 수행에 영향을 줄 수 있다.

본 연구는 이러한 임상수행평가 특히 환자진료수행을 포함한 OSCE 스테이션에 채점자가 입실하는 것과 입실하지 않은 시험장 환경이 학생들의 시험 성적에 미치는 영향과 이러한 채점자의 입실 상황에 대해 학생 응시자들이 어떻게 생각하는지에 대해 살펴보고자 한다. 이를 통해 앞으로 시행될 의사 국가시험 OSCE에서 시험장 환경과 평가방식에 대한 시사를 하려고 한다.

대상 및 방법

영남대학교 의과대학 4학년 83명을 대상으로 2005년 7월 30일 임상종합평가 일환으로 객관구조화진료시험 (objective structured clinical examination, OSCE)을 치르게 했다. 시험장을 2개로 복제하였고 시험은 환자진료를 포함한 6개 스테이션과 시뮬레이터를 이용한 임상술기 스테이션 6개, 휴식 스테이션 2개를 포함하여 총 14개 스테이션으로 구성되었다. 학생들은 출석번호 순으로 14명 1개 조로 묶어 총 5개조, 13명 1개조로 구성하여 시험을 치르게 했다. 출제자 및 평가자는 임상실습교육 및 OSCE 운영을 맡고 있는 임상교육소위원회 교수들이 담당했다. 출제자가 제출한 문항의 시나리오 및

채점 기준표는 OSCE 구성 및 진행도 (matrix)를 기반으로 작성되었으며 임상교육소위원회에서 3차례 수정 보완을 통해 최종 검증하였다. 평가는 출제자가 해당 문항의 채점을 하지 않도록 배정하여 채점하도록 했다.

각 스테이션별 시험시간은 이동시간을 포함 6분 30초였으며, 환자 진료를 포함하고 있는 6개 스테이션은 각각 10개 문항으로 구성하고 채점 척도는 예 (1점), 아니오 (0점) 2개 구간으로 구성하였다. 채점자는 이를 기반으로 한 체크리스트를 가지고서 채점하였다. 각 스테이션에 배정된 채점 교수는 각 3개 조, 41~42명의 학생들을 채점하도록 하였다. 연구를 위하여 환자 진료에 해당하는 6개 스테이션 중 소아과와 내과 2개 스테이션에 대해서 2개조 28명 학생들로 하여금 채점자가 없는 상태에서 시험을 치르게 했다. 채점자가 부재한 상태에서 시험을 치른 학생들에 대한 채점은 동일한 스테이션에 입실한 상태에서 2개조를 채점한 교수 2명이 비디오 녹화된 것을 보고 이후에 채점하였다. 내과 문제는 폐결핵 질환 진단과 환자 교육에 관한 내용으로 구성되었고, 소아과 문제는 장중첩증 진단과 질환 설명에 관한 내용으로 구성되었다. OSCE 각 스테이션 점수는 분석을 위해 100점 만점으로 환산하여 분석하였다.

환자 진료가 포함된 6개 스테이션 각각에 대한 성적은 기술통계량으로 확인하였고, 채점자 입실한 상태에서 시험을 치른 집단과 채점자가 입실하지 않은 상태에서 시험을 치른 집단간의 3학년 전체성적, 각각의 스테이션 성적, OSCE 전체성적에 대해서는 독립표본 t검증을 실시하였다. 또한 OSCE에서 채점자 입실 유무와 관련하여 채점자 교수가 부재한 상태가 환자를 진료하는 현실감을 더 주는지, 시험에 더 집중할 수 있는지, 그리고 채점자 교수가 부재한 상태를 희망하는 지에 대해서도 학생들의 의견을 묻는 설문을 실시하였고 이에 대해서도 채점자 입실 유무 집단에 따른 독립표본 t검정을 실시하였다. 이에 대한 모든 분석은 한글판 SPSS 14.0을 이용하였다.

Table 1. Total Scores of 3-year and OSCE of Raters

Total scores	Group ^{a)}	n	Mean	SD	t	p
3-year	0	28	82.91	3.80	-1.669	0.243
	1	55	84.23	3.21		

OSCE: objective structured clinical examination.

^{a)}Group 0: Group of raters not at station; Group 1, Group of raters at station.

Table 2. Scores at each OSCE Station with Standardized Patient Treatment

Station item	n	Min	Max	Mean	SD
Pediatric-Intussusception	83	20	100	73.25	18.81
Pulmonology-Pulmonary tuberculosis	83	20	90	64.94	14.76
Nephrology-Stable chronic kidney disease	83	40	100	66.99	14.54
Rheumatology-Knee pain	83	50	90	69.16	11.07
Gynecology-Lower abdominal pain	83	30	100	71.08	15.62
General surgery-Upper abdominal pain	83	70	100	87.71	8.60

결 과

1. 채점자 입실 유/무 집단 간의 3학년 성적 비교

채점자 입실 유무 집단 간에 해당 OSCE 이외에 차이 여부를 검증하기 위해 두 집단 간에 OSCE 이전의 학생들의 능력이라 할 수 있는 3학년 전체 성적과 학생들의 OSCE 전체 성적을 비교하였다. 채점자가 입실하지 않은 집단의 3학년 전체 성적은 평균 82.91점이었고, 채점자가 입실한 집단의 성적은 84.23점으로 유의한 차이가 없었다 ($p=0.243$) (Table 1).

2. 환자 진료가 포함된 스테이션 성적

OSCE 전반적인 성적 결과 중에서 환자 진료가 포함된 OSCE 6개 스테이션에 관한 성적을 분석하였다. 그 결과 소아과관련 장중첩증 증세관련 스테이션의 학생들의 평균 점수는 73.25점, 내과 폐결핵 증세 점수는 64.94점, 내과 만성 신질환 점수는 66.99점, 내과 관절통은 69.16점, 산부인과 하복부통증 71.08점, 외과 상복부통증 87.71점으로 나타났다 (Table 2).

3. 채점자 입실 유/무 집단간 스테이션 성적 차이

채점자가 입실하지 않은 상태에서 시험을 치른 집단과 채점자가 입실한 상태에서 시험을 치른 집단 간에 스테이션의 성적 차이를 확인하였다. 채점자 입실과 부재상태를 적용한 스테이션은 소아과 장중첩증 스테이션과 내과 폐결핵 스테이션 등 총 2개 스테이션으로 적용하였다. 먼저 소아과 장중첩증 스테이션에 있어서 채점자가 부재한 상태에서 시험을 치른 집단은 평균 78.57점, 채점자 입실한 상태에서 시험을 치른 집단은 70.55점으로 유의한 차이가 없었다 ($p=0.066$). 내과 폐결핵 스테이션에 있어서 채점자가 부재한 상태에서 시험을 치른 집단은 평균 66.07점, 채점자가 입실한 상태에서 시험을 치른 집단은 64.36점으로 유의한 차이가 없었다 ($p=0.621$) (Table 3).

4. 채점자 입실 유/무 집단의 OSCE 전체 성적 차이

채점자 입실 유무에 따른 OSCE 환자진료 6개 스테이션과 임상술기 6개 스테이션을 포함한 전체성

Table 3. Score Differences at each OSCE Station with Patient Treatment between Presence or Absence of Raters

Station item	Group ^{a)}	n	Mean	SD	t	p
Pediatric-Intussusception	0	28	78.57	16.27	1.865	0.066
	1	55	70.55	19.57		
Pulmonology-Pulmonary tuberculosis	0	28	66.07	14.74	0.496	0.621
	1	55	64.36	14.88		

OSCE: objective structured clinical examination.

^{a)}Group 0, Group of raters not at station; Group 1, Group of raters at station

Table 4. Total Scores of OSCE Between Presence or Absence of Raters

Total scores	Group ^{a)}	n	Mean	SD	t	p
OSCE	0	28	67.75	5.43	-1.175	0.099
	1	55	69.13	4.85		

OSCE: objective structured clinical examination.

^{a)}Group 0, Group of raters not at station; Group 1, Group of raters at station

Table 5. Opinions of Assessor's Absence at OSCE Station

Item	Group ^{a)}	n	Mean	SD	t	p
To make OSCE station more like real life	0	25	4.32	0.75	2.536	0.013
	1	52	3.85	0.78		
To concentrate better on the clinical process	0	25	4.04	0.73	2.858	0.005
	1	52	3.46	0.87		
Requesting to take OSCE without presence of assessor	0	25	3.80	0.91	3.705	0.000
	1	52	3.06	0.78		

OSCE: objective structured clinical examination.

^{a)}Group 0, Group of raters not at station; Group 1, Group of raters at station

적이 집단 간에 차이가 있는지를 확인하였다. 이를 확인한 결과 채점자가 입실하지 않은 집단은 67.75 점, 채점자가 입실한 집단은 69.13점으로 유의한 차이가 없었다 (p=0.009) (Table 4).

5. OSCE 채점자 입실 시험환경에 대한 집단간 의견 차이

OSCE 실시 직후 응시 학생들에게 채점자 입실

유무와 관련하여 5구간 척도 (1구간 1점)의 설문을 조사하였다. 채점자가 없는 상황에서 시험을 치르는 것이 환자를 진료하는 현실감을 더 느끼게 하는지에 대해서 채점자가 부재한 상태를 경험한 집단은 4.32점, 채점자가 입실한 상황에서만 시험을 치른 집단은 3.85점으로 차이가 있었다 (p=0.014). 채점자가 없는 상태가 시험을 치르는 데에 더 집중할 수 있게 하는지에 대해서 채점자가 부재한 상태를 경

험한 집단은 4.04점, 채점자가 입실한 상황에서만 시험을 치른 집단은 3.46점으로 차이가 있었다 ($p=0.005$). OSCE에서 스테이션에 채점자가 입실하지 않은 상태에서 시험을 치르는 것을 희망하는 지에 대해서 채점자가 부재한 상태에서 시험을 치른 경험이 있는 집단은 3.80점, 채점자가 입실한 상황에서만 시험을 치른 집단은 3.06점으로 차이가 있었다 ($p=0.000$) (Table 5).

고 찰

OSCE는 전통적인 지필중심의 평가보다 임상교육의 목적에 보다 부합되는 것이 사실이다 (Probert *et al.*, 2003). 이러한 OSCE에 합격선이 적절히 결정되어 있지 않은 경우 평가적 측면에서의 대외적 영향력은 미흡할 수밖에 없다 (Morrison *et al.*, 1996). 그에 따라 OSCE는 적절한 준비와 처리 절차가 동반되지 않으면 평가의 신뢰도와 타당도를 확보하기 어렵다 (Newble & Dawson, 1994). 이는 Norman (1989)이 지적한 바와 같이 실제적인 임상수행을 평가하는 것이 실질적인 상황을 기반으로 매우 복합적인 내용을 측정해야 하기 때문에 임상 수행에 관한 인지능력을 평가하는 것이 통계학적으로나 측정학적으로도 훨씬 유효하지 못할 수도 있기 때문이다.

지금까지 OSCE 운영에서 흔히 간과되는 점은 시험장 환경을 실제 임상진료 상황과 동일하게 해야 한다는 점이다. 스테이션 단위로 이루어지는 OSCE에서 신뢰도 검증은 주로 채점자내, 채점자간, 시험장 복제 및 표준화환자에 의한 채점의 신뢰도 검증 등 채점 자체에 대해서는 많은 관심을 기울였지만 시험장 환경 자체에 대해서는 깊이 고려되지 못했다. 본 연구는 OSCE에서 시험장이라고 할 수 있는 스테이션에 채점자가 입실한 상태가 응시 학생들에게 영향을 미치는지를 확인하였다.

본 연구에 따르면 채점자가 입실하지 않은 집단의 평균 점수가 채점자가 입실한 상태에서 시험을 치른 집단보다 평균점수가 높게 나타났지만 통계학적으로 유의하지 않은 것으로 나타났다. 그럼에도 채점자가 입실하지 않은 상태에서 시험을 치른 학

생들은 환자진료가 포함된 OSCE에서 채점자가 입실하지 않는 것을 더욱 바람직한 것으로 응답했으며 이는 채점자가 입실한 상태에서만 시험을 치른 집단과 통계적으로 유의미한 차이가 있는 것으로 나타났다. 이러한 점에서 보면 OSCE 시험장 환경에서 채점자 입실은 고려해야 할 점이다. 특히 10분 내외 시간동안 환자면담과 신체진찰을 통해 진단과 환자교육까지 해야 할 상황에서 학생들의 긴장감은 매우 높을 것이며 채점자의 입실은 그에 대한 긴장을 더 높일 가능성이 있다. 그에 따라 채점자가 입실하지 않은 상태에서 OSCE 운영 방안을 고려할 필요가 있다.

비디오 녹화를 통한 방법은 채점자가 입실하지 않은 상태에서 채점할 수 있는 좋은 방법중의 하나이다. Vivekananda-Schmidt *et al.* (2007)에 의하면 비디오 녹화를 통한 채점은 채점자 동원이나 서로 다른 응시자 집단간의 차이를 줄일 수 있는 점 등 효과적인 측면이 많다. 특히 녹화된 내용을 가지고서 훈련받은 표준화 환자나 교수가 채점할 경우 OSCE 상황에 대해 간과하거나 실제 상황에서 놓칠 수 있는 부분을 재확인 할 수 있다는 점에서 채점의 정확성을 높일 수 있을 것이다.

이 연구는 채점자가 입실하지 않은 상태에서 OSCE를 치른 집단과 채점자가 입실하지 않은 상태에서 OSCE를 치른 집단간의 차이를 확인하였다. 하지만 채점자가 입실하지 않은 경우를 모의환자진료가 포함된 일부 경우에 적용했다. 본 연구의 결과를 일반화하는 데에는 아직 한계점이 있다. 이와 관련하여 후속 연구에 대한 제언을 하자면 다음과 같다. 먼저 채점자가 입실하지 않은 스테이션의 수를 확대시켜 적용하여 검증하는 연구가 필요할 것이다. 이러한 과정에서 녹화 영상을 통해 표준화환자의 채점과 교수 채점자에 의한 채점을 비교하여 일치도와 정확도를 검증하는 것은 표준화환자에 의한 채점의 신뢰도 범위를 보다 명확히 할 수 있을 것이다. 다음으로 채점자가 실시간 채점하는 것과 비디오 녹화하여 채점하는 것과의 비교를 통하여 차이를 검증하는 연구도 필요하다. 앞서도 언급한 바와 같이 실시간 채점은 학생 응시자들의 수행을

간과할 수 있을 뿐만 아니라 많은 연구에서도 지적 하듯이 장시간 OSCE 참여에 의한 피로로 인한 채점자의 객관성을 떨어뜨릴 수도 있다. 그에 따라 이러한 검증은 채점 방식 운영에 대한 또 다른 시사할 것이다.

ACKNOWLEDGEMENTS

This work was supported by the Research Fund from the Korean Society of Medical Education (2005).

REFERENCES

- Harden, R.M., & Gleeson, F.A.(1979). Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ*, 13, 41-54.
- Hodges, B.(2003a). Validity and the OSCE. *Med Teach*, 25, 250-254.
- Hodges, B.(2003b). OSCE! Variation on a theme by Harden. *Med Educ*, 37, 1134-1140.
- Kim, K.H.(2000). A psychometric approaches of the validity verifications in performance assessments. Doctorial dissertation, The Graduate School, Ewha University.
- Kirby, R.L., & Curry, L.(1982). Introduction of an objective structured clinical examination (OSCE) to an undergraduate clinical skills programme. *Med Educ*, 16, 362-364.
- Morrison, H., McNally, H., Wylie, C., McFaul, P., & Thompson, W.(1996). The passing score in the objective structured clinical examination. *Med Educ*, 30, 345-348.
- Newble, D., & Dawson, B.(1994). Guidelines for assessing clinical competence. *Teach Learn Med*, 6, 213-220.
- Newble, D.(2004). Techniques for measuring clinical competence: objective structured clinical examinations. *Med Educ*, 38, 199-203.
- Norman, G.R.(1989). Reliability and construct validity of some cognitive measures of clinical reasoning. *Teach Learn Med*, 1, 194-199.
- Park, H., Lee, J., KIM, S., Kim, K., & Park, H.(1999). The effect of using two duplicated examination sites to simulate the same cases in the OSCE reliability. *Korean J Med Educ*, 11, 37-52.
- Park, H., Lee, J., Hwang, H., Lee, J., Choi, Y., Kim, H., & Ahn, D.H.(2003). The agreement of checklist recordings between faculties and standardized patients in an objective structured clinical examination (OSCE). *Korean J Med Educ*, 15, 143-152.
- Probert, C.S., Cahill, D.J., McCann, G.L., & Ben-Shlomo, Y.(2003). Traditional finals and OSCEs in predicting consultant and self-reported clinical skills of PRHOs: a pilot study. *Med Educ*, 37, 597-602.
- Regehr, G., MacRae, H., Reznick, R.K., & Szalay, D.(1998). Comparing the psychometric properties of checklist and global rating scales for assessing performance on an OSCE-format examination. *Acad Med*, 73, 993-997.
- Shin, J., Lee, S., & Park, H.(2005). Standardized patients' accuracy in recording checklist items during clinical performance examinations. *Korean J Med Educ*, 17, 197-203.
- Vivekananda-Schmidt, P., Lewis, M., Coady, D., Morley, C., Kay, L., Walker, D., & Hassell, A.B.(2007). Exploring the use of videotaped objective structured clinical examination in the assessment of joint examination skills of medical students. *Arthritis Rheum*, 57, 869-876.