

의학교육평가로서 전문의자격시험 실기시험의 국내외 현황과 발전방안

조 정 진¹ · 노 혜 미¹ · 김 승 호² · 권 호³ · 박 영 민⁴ · 최 병 민⁵ | ¹한림대학교 의과대학 가정의학과교실, ²연세대학교 의과대학 응급의학교실, ³가톨릭대학교 의과대학 성형외과교실, ⁴가톨릭대학교 의과대학 피부과교실, ⁵고려대학교 의과대학 소아과학교실

The current status and development of a skill examination for the Korean specialty certification examination

Jung Jin Cho, MD¹ · Hye Mi Noh, MD¹ · Seung Ho Kim, MD² · Ho Kwon, MD³ · Young Min Park, MD⁴ · Byung Min Choi, MD⁵

¹Department of Family Medicine, Hallym University College of Medicine, Chuncheon, ²Department of Emergency Medicine, Yonsei University College of Medicine, Seoul, ³Department of Plastic and Reconstructive Surgery, The Catholic University of Korea College of Medicine, Uijongbu, ⁴Department of Dermatology, The Catholic University of Korea College of Medicine, Seoul, ⁵Department of Pediatrics, Korea University College of Medicine, Ansan, Korea

The Specialty Certification Examination is an important part of the assessment of medical education. However, the step 2 skill examinations of 26 medical specialties in Korea are insufficient with respect to achieving the objective of practical examinations that evaluate clinical skill and competence. Among the current step 2 skill examination methods, picture testing using slides or reading of pathology slides/radiologic images is more suitable for testing cognition and knowledge than for testing performance. The oral examination has low reliability because of its relatively short testing period and absence of scoring criteria. In addition, the Specialty Certification Examination is a high-stakes test and the performance during the training course is not reflected in the skill examination. We have reviewed the various skill examinations including clinical practice examinations, objective structured clinical examinations of the United States and Canada, and work-based assessments of the United Kingdom. Based on the review, we suggest some plans for improving the Korean Specialty Certification Examination.

Key Words: Specialty Certification Examination; Performance assessment; Skill examination; Work based assessments

서론

수련병원별 수련의 질 차이가 크기 때문에 전문의자격시험은 졸업 후 의학교육에서 중요한 역할을 수행하고 있다.

현재 각 전문학회에서 시행하고 있는 2차시험은 빔프로젝트로 사진이나 그림을 제시하는 슬라이드시험으로 시행되는 사례가 많다. 사진을 이용한 슬라이드시험은 의학교육 평가 측면에서 실기시험이라기 보다 사진이나 그림을 슬라이드로 제시할 뿐이므로 지면에 사진을 제시하는 필기시험과 같은 인지영역 평가시험이다. 따라서 진정한 의미의 실기시험으로 개선이 필요하다. 이 글에서는 의학교육평가로서 전문의자격시험의 의미를 조명해보고, 외국 전문학회에서 시행 중인 실기시험 현황을 파악해보고자 한다. 또한 26개 전문학회를 대상으로 2차실기시험에 대한 인식도 및 의견을 설문조사를 통해 수렴하여, 국내에 적용 가능한 실기시험 개선 방

Received: March 13, 2014 Accepted: March 27, 2014

Corresponding author: Jung Jin Cho
E-mail: fmjicho@naver.com

© Korean Medical Association

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.

향에 대한 대안을 제시해 보고자 한다. 이 논문은 조정진 등이 2013년 대한의사협회 연구용역보고서로 제출한 '전문의 자격시험 실기시험 개선방안'을 토대로 요약 정리한 글이다.

의학교육평가로서의 전문의자격시험

교육평가란 이미 정하여진 교육목표를 얼마나 달성하였는가 즉 교육목표와 성취 사이의 일치도를 알아내기 위하여 정보를 수집, 분석, 해석하여 결정하는 체계적인 과정이다. 전문의자격시험은 의학교육평가의 한 영역으로 전공의가 수련교과과정을 이수하고 해당과 전문의로서의 임상역량이 있는가를 총괄적으로 평가하는 시험이다. 임상역량은 다양한 정의가 있지만 미국의 전공의 수련교육인증위원회(Accreditation Council for Graduate Medical Education)는 환자진료, 의학적 지식, 진료에 기초한 학습과 향상, 대인 및 의사소통기술, 계통적 진료 등 5가지 영역을 전문의가 되기 위한 임상역량으로 규정하고 있다.

1. 의학평가 단계와 평가도구

Crossley 등[1]은 의학교육평가를 교육 목적에 따라 인지영역(cognition), 기술영역(skill)과 태도영역(attitude)으로 분류하였다. Miller [2]는 인지영역과 행동영역(behaviors: performance, action)으로 의학교육 평가단계를 구분하였다.

국내 전문의자격시험도 이러한 의학교육 평가단계를 반영하여 '전문의의 수련 및 자격 인정 등에 관한 규정 시행규칙' 제13조 2항에 의하면 1차시험은 필기시험, 2차시험은 실기시험 또는 수술시험이라고 법으로 규정되어 있다. 기술영역과 태도영역 또는 행동영역을 평가하는 실기시험에 적용할 수 있는 평가도구는 2가지 영역으로 나눌 수 있는데, 수행평가영역(performance: assessment of 'shows how')과 실행평가영역(action: assessment of 'do')이다[3].

1) 수행평가 영역

의학교육평가의 실기시험 중 최근 'show how'를 보는, 즉, 행동영역의 수행(performance)을 평가하는 대표적인 평가도구로는 객관구조화진료시험(objective structured clinical

examination, OSCE)이나 임상진료시험(clinical practice examination, CPX) 등이 있다. Long & short cases 같은 표준화 환자를 이용하지 않은 진짜 환자를 이용한 사례 평가는 OSCE 도입 이전에 많이 활용되었다. 이 평가방법은 선정된 환자를 30-45분간 평가자 없이 진료한 후 평가자에게 진료내용을 보고하여 평가받는 것이다. 최근에는 사례 표본 선정과정에서 내용 특이도, 신뢰도 및 긴 시험시간의 문제로 표준화 환자를 이용한 OSCE로 대체되고 있다.

OSCE는 수험생은 비교적 짧은 5-15분으로 구성된 스테이션에서 특정 술기의 수행능력을 단편적으로 확인하는 과제를 수행하는 시험양식이다. CPX는 비교적 긴 10-12분 동안 시행되는데, 수험생은 표준화 환자를 대상으로 실제 환자처럼 병력청취를 시행하고, 신체검사를 수행하며, 임상검사를 내고, 진단을 내리며 치료계획을 세우고, 상담을 수행하여 포괄적인 임상수행능력을 평가하는 평가방법이다. CPX를 OSCE의 한 형태로 분류하거나 이에 다른 시험으로 구분하기도 한다[4].

2) 실행평가 영역

Miller [2]의 평가 단계 최상위 영역인 행동영역의 실행영역을 보는 실기시험으로는 Mini Clinical Evaluation Exercise, Direct Observation of Procedural Skills, Clinical Work Sampling, Logbooks, Portfolios, Consultation Observation Tool, Case-based Discussion (CbD), Multi-Source Feedback, Patient Satisfaction Questionnaire (PSQ), Clinical Supervisors Report (CSR) learning 등이 있다. Mini Clinical Evaluation Exercise (Mini-CEX)는 평가자 관찰 아래 실제 진료한 환자를 이용하여 초진 진료를 20-30분 시행하여 양식화된 체크리스트로 평가한다. Clinical Work Sampling은 실제 일하는 환경에서 수행능력을 평가하는 것으로 표본 사례를 선정하여 초진환자 진료, 입원, 퇴원, 병동 경과관찰 등의 정보를 수집하여 평가한다. Log와 Portfolios은 즉 실제 진료한 사례나 증례를 제시하고 평가 받는 것으로 상호평가가 가능하다[5]. 미국 전문의자격시험에서 주로 시행하는 진료중심 수술시험(practice based oral examination)은 필기시험 합격 후 전문의가 진료한 사례 목록 즉 case logs를 제출하고, 이 사례 목록에서 일부

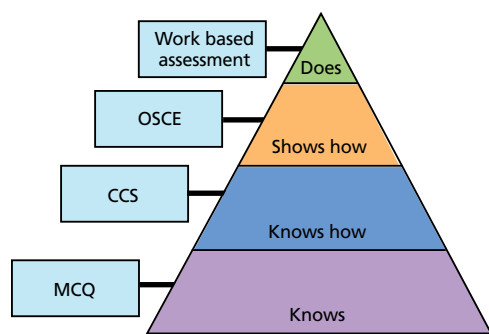


Figure 1. The examination methods and Miller's pyramid of competence. OSCE, objective structured clinical examination; CCS, computerized case simulation; MCQ, multiple choice questions.

사례를 선정하여 증례토론 구술시험을 시행하고 있다. 전문 의자격시험에서 실제 실행한 진료에 연계하는 방식은 2가지 형태가 가능한데, 하나는 전공의 수련과정에서의 실제 시행한 진료내용을 평가하는 것을 직무중심평가(work based assessment, WBA) 라고 하며[6], 다른 하나는 필기시험 합격 후 전문의 과정 중 진료목록을 제출하여 평가받는 것으로 진료중심시험(practice based examination, PBA)이라고 할 수 있다(Figure 1).

2. 교육평가원칙과 평가도구별 특성

의학교육 평가인 전문의자격시험도 평가목표를 달성하기 위해서는 타당도, 신뢰도, 객관도, 실용도 같은 교육평가의 기본적인 요건을 충족하는 것이 필요하다.

타당도란 평가도구 내용의 충실성으로 무엇을 측정하려고 하느냐의 개념이다. 즉 해당과 전문의에게 필요한 내용인지 또는 출제 분야가 수련교과과정에 포함되고 해당과 전문의가 알아야 할 영역인지를 살펴보는 것이다. 타당도에는 내용타당도, 안면타당도, 구성개념타당도, 준거타당도가 있다 [7]. 신뢰도란 얼마나 안정성을 갖고 일관성 있게 측정 또는 평가하고 있는지의 문제로서 어떻게 측정하느냐의 개념이라고 할 수 있다. 객관도는 채점자에 의해 결정되는 일종의 신뢰도로 동일한 답안지를 채점할 때 마다 일정하게 동일한 점수가 나오느냐는 것이다. 실용도란 시간, 비용, 노력이 얼마나 현실적이고 실용적이나 하는 개념이다.

신뢰도는 시험유형에 따라 다르며, 시험유형이 같더라도 시험시간에 따라 달라지는데, 다음은 Table 1은 유형별 시

Table 1. Reliability according to examination methods and testing times

Examination methods	Testing times (hr)			
	1	2	4	8
Multiple choice ^{a)}	0.62	0.76	0.93	0.93
Patient management problem ^{a)}	0.36	0.53	0.69	0.82
Key feature case (write-in) ^{a)}	0.32	0.49	0.66	0.79
Oral examination ^{b)}	0.50	0.69	0.82	0.90
Long case examination ^{b)}	0.60	0.75	0.86	0.90
Objective structured clinical examination ^{a)}	0.54	0.69	0.82	0.90
Mini-clinical exercise ^{c)}	0.73	0.84	0.92	0.96
Practice video assessment ^{b)}	0.62	0.76	0.93	0.93
Incognito standardised patients ^{c)}	0.61	0.76	0.82	0.86

From Van der Vleuten CP, et al. Med Educ 2005;39:309-317, with permission from John Wiley and Sons [8].

^{a)} One-facet all random design with items crossed with persons.

^{b)} Two-facet all random design with judges (examiners) nested within items within persons.

^{c)} One-facet all random design with items nested within persons.

험시간에 따른 신뢰도에 대한 대표적인 연구결과이다[8]. 구술시험 특히 짧은 구술시험은 다른 시험에 비해 신뢰도가 낮은 것이 문제로 지적되고 있어 최근 고부담시험(high skate test)인 자격시험에서 배제되고 있는 경향이 있다. 다만 12-16개의 사례를 활용한 구술시험(long cases)의 신뢰도는 0.8로 높아진다는 보고가 있다. OSCE의 신뢰도가 높아 실기 시험 중 좀 더 선호되고 있다.

교육평가의 목적에 따라 교육평가 원칙인 타당도와 신뢰도의 중요도가 달라진다. 교육평가를 시기로 구분하면 진단평가, 형성평가, 총괄평가로 분류한다. 형성 평가는 교육과정 도중에 이루어지며, 목표를 향해 올바른 방향으로 가고 있는가를 점검하는 과정이라고 할 수 있다. 즉 평가 결과에서 잘못된 점, 부족한 점을 보완시켜 올바른 궤도로 바로잡는 목적과 기능을 가졌을 때 그것을 형성 평가라고 하며 이에 따라 학습목표, 학습방법의 개선을 할 수 있는 되먹이 기능이 있음으로 교육적으로 매우 큰 의의가 있다. 반면 총괄평가는 일정한 과정이 끝난 뒤에 소기의 목표에 도달되어 있는가를 최종적으로 알아보는 판정의 기능과 경쟁목적을 갖는 평가를 총괄하여 평가한다. 진급, 수료, 졸업 등의 가부 여부를 결정 내리는 경우이다. 전문의자격시험을 평가하기로 구분하면 총괄평가이다. 총괄평가도 평가방법에 있어 순위에 역점을 둔 규준평가와 최소한의 자격기준에 역점을 둔 준거 평가로 구분할 수 있다. 졸업시험은 대개 규준평가이고

Table 2. Assessment methods in surgical training in the United Kingdom

Assessment methods	Description	Applications
Written tests		
Other	MCQs, EMQs, and SAQs are well established written exam formats used for multiple purposes	Selection to a programme of training certification
GCSEs	Exams on selected subjects depending on the chosen field of future higher education	Selection to medical school
Application form	Depending on the purpose they combine different elements, such as a CV and short essay type questions	Selection to a programme of training Selection to medical school
Amplitude tests (UKCAT, BMAT)	Measure performance across a range of mental abilities	Selection to medical school
Orals and vivas		
MRCS FRCS	Assessment of knowledge by face-to-face interaction between the assessor and the candidate	Certification
Observational assessments of clinical competence		
Work based assessments (DOPS, Mini-CEX, CBD, PBA, OSATS)	Assessment of clinical competence with direct observation of the candidate performance in a real clinical or simulated context	Progression in a programme of training
OSCE		Certification
Interviews	Depending on the purpose they combine different elements, such as skills stations, CV stations, and assessment centres	Selection to a programme of training Selection to medical school
Portfolios	Record of achievements and experiences	Progression in a programme of training

From Evgeniou E, et al. J Educ Eval Health Prof 2013;10:2, according to the Creative Commons license [9]. MCQ, multiple choice questions; EMQ, extended matching question; SAQ, short answer question; GCSE, General Certificate of Secondary Education; CV, curriculum vitae; UKCAT, UK Clinical Aptitude Test; BMAT, Biomedical Admissions Test; MRCS, Member of the Royal College of Surgeons; FRCS, Fellow of the Royal College of Surgeons; DOPS, direct observation of procedural skills; Mini-CEX, Mini-clinical evaluation exercise; CBD, case-based discussion; PBA, procedure-based discussion; OSATS, objective structured assessment of technical skills; OSCE, objective structured clinical examination.

전문의고시는 준거평가이다.

형성평가에는 타당도가 좀 더 주안점이 주어지며, 총괄평가에서는 신뢰도에 좀 더 초점을 둔다. 평가도구별 특성에 따라 타당도와 신뢰도가 달라지기 때문에 시험의 목적에 따라 평가도구의 활용도가 달라진다. 인지영역을 주로 평가하는 필기시험은 신뢰도, 객관성, 실용도가 높은 시험이다. 실기시험은 각 전문학회의 전문의로서의 전문적 진료능력을 평가하는데 있어 단순한 인지영역 평가인 필기시험이 평가할 수 없는 태도 및 기술, 수행, 행동을 평가할 수 있는 중요한 도구이다. 실기시험은 타당도가 높은 반면, 신뢰도가 상대적으로 낮다. 따라서 필기시험은 총괄평가에 더욱 적합하고, 실기시험은 형성평가에 적합하다고 할 수 있다. 영국의 외과 전문의시험에서는 WBA는 수련과정의 타당성을 평

가하는 형성평가로, 총괄평가인 최종 자격시험에는 신뢰도가 높은 필기시험과 OSCE를 활용하는 것은 평가도구의 특성을 잘 반영한 것으로 보여진다[9,10] (Table 2).

외국의 전문의자격시험 실기시험 현황

1. 미국 전문의자격시험

미국은 우리나라와 달리, 정부에서 전문의 자격을 부여하는 것이 아니라, 전문학회들이 연합단체를 구성하여 전문의시험을 자율적으로 관장하고 있다[11]. 미국 전문의자격과 관련된 단체는 미국전문과별위원회(The American Board of Medical Specialties, ABMS), 미국전문의협회(The American Board of Physician Specialties, ABPS) 등이 있다. ABMS는 1933년에 설립된 비영리기관으로 전문의자

격에 관한 기준 제정 및 평가와, 자격부여를 위한 전문의 인증 및 교육을 개발해왔으며, 24개로 구성된 전문과학회(Specialty Member Boards)가 150여 개의 전문의 및 세부 전문의 자격을 관리한다[12]. 각 회원단체는 전문의자격을 요구하는 사람들이 교육기준에 따른 전공의 훈련을 이수하였는지 점검하고, 시험을 통해 검증하여, 전문의자격을 부여한다. ABMS 전문의자격 요건은 다음과 같다. 4년제 대학을 졸업 후, 인증 의과대학에서 의사학위를 취득하고, 인증 수련기관에서 3-5년의 수련을 수료한 진료면허 소지자이다. 자격요건을 갖춘 후, ABMS에서 관장하는 필기시험과 일부 과의 경우 구두시험에 합격하면, 전문의자격을 부여 받는다. 그리고 ABMS에서는 2006년부터 자격을 받은 전문의들의 질 관리를 위해, 자격면허 갱신과정

Table 3. The skill examination of American Boards of Specialties and Royal College of Physicians and Surgeons of Canada

		Examination methods		
		Oral exam	Practical exam	OSCE/CPX
America	ABMS	Anesthesiology, radiation oncology, plastic Surgery, ophthalmology, radiology, general surgery, obstetrics and gynecology, emergency medicine, orthopedic surgery		
	ABPS	Family medicine obstetrics, anesthesiology, radiation oncology, plastic surgery, ophthalmology, radiology, general surgery, obstetrics and gynecology, emergency medicine		
Canada	CFPC			Family medicine
	RCPSC	Anesthesiology, dermatology, pathology, radiology, emergency medicine, general surgery, neurosurgery, internal medicine, hematology	Pathology, cardiology	Radiology, neurology, ophthalmology, obstetrics and gynecology, cardiology, gastroenterology, rheumatology

OSCE, objective structured clinical examination; CPX, clinical practice examination; ABMS, The American Board of Medical Specialties; ABPS, The American Board of Physician Specialties; CFPC, The College of Family Physicians of Canada; RCPSC, The Royal College of Physicians and Surgeons of Canada.

(Maintenance of Certification, MOC)를 도입하여 운영 중이다. 현재까지 ABMS에서 80만 여명의 의사에게 자격을 부여하였으며, 이중 45만여 명이 MOC 과정에 참여하고 있다. ABMS 이외에, 미국의 전문의자격을 관리하는 기관으로는 ABPS가 있다. ABPS는 미국전문의학협회(American Association of Physician Specialists)의 산하 기관으로, 12개의 운영위원회가 17개의 전문의 및 세부 전문의 자격을 관리한다[13]. 전문의 자격 요건 및 시험 과정은 ABMS의 규정과 유사하며, 기초지식, 전문기술 및 지속적인 교육과 훈련에 대한 평가를 시행하고, 전문의자격시험은 필기시험과 일부 과의 경우 구두시험도 시행한다. 또한 ABMS와 마찬가지로, 면허 갱신과정을 운영하고 있다. 미국 전문의자격시험은 필기시험과 구술시험으로 이루어져있으며, 캐나다와는 달리 임상진료시험은 따로 시행하고 있지 않다. 필기시험은 모든 과에서 시행하고 있으며, 구술시험은 각 전문과의 특성에 따라, 시행 여부가 다르다. 구술시험을 시행하는 학회는 ABMS의 소속 학회 중에서는 안과, 정형외과, 성형외과, 마취과, 산부인과, 외과, 영상의학과, 방사선종양학과, 응급의학과 학회, ABPS의 소속 학회로는 가정

의학-산과, 안과, 성형외과, 마취과, 산부인과, 외과, 영상의학과, 방사선종양학과, 응급의학과 학회이다(Table 3).

구술시험은 각 전문학회별로 다양한 형식과 문항 수로 구성하여 자율적으로 시행하고 있다. 대체로 시험관이 임상중례를 제시하고, 질환의 진단과 치료방법에 답하는 형식으로 이루어지는데, 정형외과, 성형외과, 산부인과 등 일부 과의 경우에는 진료중심 구술시험(practice based oral exam)을 시행한다. 즉, 응시자가 수련기간에 경험했던 환자 명단을 학회에 제출하면(case logs), 그 사례들 중, 일부 사례를 선정하여 시험관이 질문하고 토론하는 형태로 이루어진다. 응급의학과 구술시험의 경우, 한 명의 환자로 이루어진 중례(single patient

encounter, SPE)와 실제 응급실에서의 상황에서처럼, 여러 명의 환자로 이루어진 모의 중례(simulated situation encounter, SSE)의 두 종류가 있다. SPE는 감독관이 모의 환자 역할을 하고(simulated clinical encounters), 환자의 기초 데이터를 제공하면, 수험자가 병력청취, 신체진찰, 검사 처방 등을 시행하여, 신체진찰 소견이나 검사결과 등 추가적인 데이터를 얻어내서, 환자를 치료해나가는 형식이라는 점이 CPX와 유사하지만, 표준화환자가 아닌 감독관이 환자 역할을 한다는 차이점이 있다. SSE는 수험자가 첫 번째 환자를 다 치료하기 전에, 감독관이 두 번째와 세 번째 환자에 대한 자료를 제시하여, 여러 환자들의 중증도를 고려하여, 적절한 치료를 잘 수행하는지를 평가한다.

2. 캐나다 전문의자격시험

캐나다 전문의자격시험은 캐나다 가정의학회(The College of Family Physicians of Canada, CFPC)와 캐나다 왕립전문의학회(The Royal College of Physicians and Surgeons of Canada, RCPSC)에서 관장한다.

전문의자격시험은 필기시험과 실기시험으로 이루어지며,

실기시험은 각 전문과의 특성에 따라 실습시험, 구술 시험, OSCE 중 선택되며, 전문과별로 사례 수, 시험시간, 평가항목은 각각 다르다. 실습시험은 병리과, 심장내과에서, 구술 시험은 피부과, 일반외과, 신경외과, 흉부외과, 소아외과, 혈관외과, 마취과, 응급의학과 등에서, OSCE/CPX는 일반내과, 류마티스내과, 소화기내과, 신경과, 가정의학과 등에서 시행된다(Table 3).

CFPC는 가정의학과 전문의 수련 과정 기준 제정 및 평가, 전문의 자격 인증, 지속적인 교육을 시행한다[14]. 전문의 자격 요건은 CFPC에서 인증한 가정의학과 수련과정을 적어도 24개월 이상 받고, 캐나다의사자격시험(Medical Council of Canada Qualifying Examination Part 1)을 마친 자로, 수련과정이 6개월 남은 시점부터 전문의자격시험에 응시할 수 있다. CFPC의 자격시험은 필기시험과 실기시험으로 구성되어 있으며, 필기시험은 의학지식과 문제해결 기술을 평가하는 단답형 문제 해결형 문항으로 이루어지며, 6시간 동안 시행된다. 실기시험은 OSCE와 모의 외래진료 구술시험(simulated office orals, SOO)로 구성되어 있다. OSCE는 표준화 환자를 대상으로 임상상황에 대한 문제가 주어지고, 병력청취 및 신체진찰 능력과 면담 및 대화기술을 평가하며, 총 8개 문항으로 각 10분씩 시행된다. SOO는 감독관이 환자 역할을 하고, 가정의과 진료 시 처할 수 있는 상황을 최대한 재현하여, 건강문제의 정의와 관리, 환자 중심적인 임상 방법과 좋은 환자-의사 관계를 형성하는 지 등에 초점을 맞추어 평가하며, 정확한 진단을 내리는 것에 대한 배점은 높지 않다. 총 4문항으로 각 15분씩 시행된다.

가정의학과 이외에 다른 전문과의 자격시험은 RCPSC에서 관장하며, 전문의자격시험 자격 조건은 RCPSC에서 인증한 전문과목 수련 과정을 마친 자로, 미국이나 캐나다 이외에, RCPSC에서 인증하는 29개국의 수련 과정을 마친 외국인도 자격시험에 응시할 수 있다[15]. 실습시험은 슬라이드나 그래픽에 대해 판단할 수 있는 능력을 평가한다. 해부병리, 일반병리, 혈액종양병리, 신경병리, 미생물학에서의 슬라이드 시험이나, 심장내과에서 심전도, 흉부촬영사진, Holter monitoring 결과, 운동부하 심전도 검사, 심장 초음파, 조영술 등에 대한 사진이나 간단한 동영상에 주어지고,

이에 대한 문제를 풀게 된다.

구술시험은 주로 외과 계열과 표준화 환자를 대상으로 시험을 시행할 수 없는 일부 과에서 시행된다. 외과에서는 수술방법, 수술적 문제해결, 판단력, 의사소통기술, 기초지식 등에 대한 구술시험이 2-3시간에 걸쳐 진행된다. 응급의학과는 2시간 동안 6개의 방에서, 흉부촬영사진, 심전도, 검사실 결과, 사진 등을 이용한 6개의 시나리오로 1:1 구술시험을 시행하게 되며, 응급의학 전반에 대한 지식 및 질환의 병태생리학적 기초 등을 평가한다. 호흡기내과는 각 1시간 정도 2세션에서 흉부촬영사진, 폐기능 검사, 컴퓨터단층촬영, 검사실 검사 등을 이용한 시나리오를 제시하여, 수험자의 임상능력을 평가한다.

OSCE는 표준화 환자를 대상으로 임상적 기술을 평가하는 시험이다. 일반 내과는 8개의 station에서 2시간 가량 시험이 이루어지며, 외래환자, 입원환자, 중환자실 등의 상황에서의 임상능력을 평가하고자 한다. 임상상황에서 신체진찰, 임상적 문제해결에 대한 접근, 진단적 검사 적용 및 검사 결과의 해석, 동료나 환자 및 보호자들과의 의사소통 능력, 윤리적 문제를 다루는 능력 등이 평가대상이 된다. 신경과나 류마티스내과는 10명의 표준화 환자를 대상으로 각 15분간 시험이 진행된다. 소화기내과는 표준화 환자를 대상으로 2시간 가량 시험을 치르며, 검사실 검사, 내시경 및 병리학적 소견, 영상검사 등의 진단적 검사를 해석하고, 신체진찰 기술, 태도, 윤리적 측면, 의학적 문제의 해결 등의 측면을 평가한다.

3. 영국 전문의자격시험

영국의 전문의자격시험은 영국왕립일반차의료전담의학회(Membership of the Royal College of General Practitioners, MRCGP), 영국왕립의사학회(Membership of the Royal Colleges of Physicians of the United Kingdom, MRCP UK), 영국왕립외과학회(Membership of the Royal College of Surgeons, MRCS)가 관장한다.

일반차의료전문의(가정의) 자격은 의사면허 취득 후, 기초수련의 과정 2년을 마친 후 일반차의료 전문의 수련프로그램을 마치고, 수련 후 통과시험(Certificate of Completion of

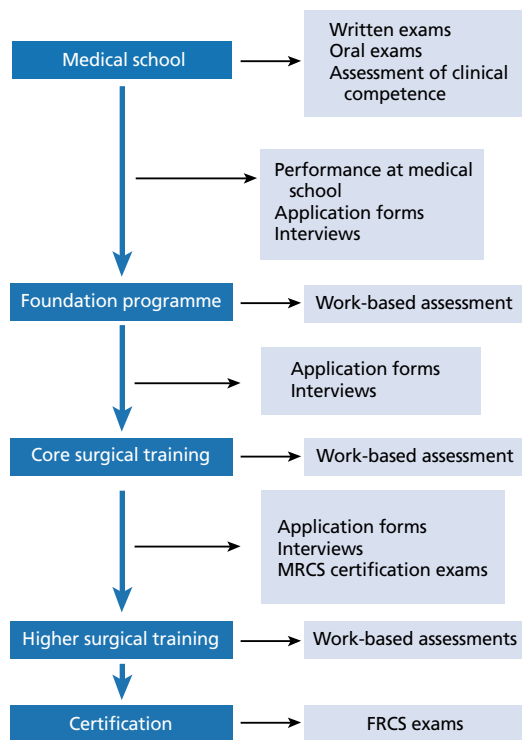


Figure 2. The surgical training pathway. MRCS, Membership of the Royal College of Surgeons; FRCS, Fellow of the Royal College of Surgeons. From Evgeniou E, et al. J Educ Eval Health Prof 2013;10:2, according to the Creative Commons license [9].

Training, CCT)을 통과한 후 MRCP 시험에 응시할 수 있다. MRCP는 필기시험(Applied Knowledge Test, AKT), 임상진료 실기시험(Clinical Skills Assessment, CSA), WBA의 세 가지 항목으로 구성된다. AKT는 전공의가 독립적인 일차의료를 할 수 있는 지식수준을 가지고 있는가를 평가한다. Clinical medicine 160문항, evidence interpretation 20문항, organisational 20문항으로 구성되며, 연중 1월, 4-5월, 10월 3회 시험이 시행된다. CSA는 일차진료에 적절한 임상적, 전문적 의사소통 능력 및 기술을 가지고, 통합할 수 있는 능력이 있는지를 평가한다. 연중 2-3월, 5월, 11월 3회 시험이 시행된다. WBA는 Mini CEX, CbD, PSQ, CSR 등 다양한 평가도구를 이용하며, 수련 현장에서 전문 역량부문에서 전공의의 성장 과정을 평가한다[16].

일차의료 전문의 이외에 의뢰전문의는 의사면허 취득 후, 기초 수련의 과정 2년을 마친 후 전문과 선택하여 수련을 받고 CCT를 통과한 후 MRCP UK 또는 MRCS에 소속된 해당 전문학회에서 시행하는 자격시험을 치른다. 의뢰전문의

자격시험을 MRCS를 예시로 살펴보면 Figure 2와 같다[9]. MRCS 자격시험은 필기시험과 OSCE실기시험으로 이루어졌으며, 이외에 수련과정에 대한 WBA평가가 포함되어 있다.

국내 전문의자격시험 실기시험 현황과 전문학회 인식도 조사

2013년 대한의사협회 연구용역 보고서인 ‘전문의자격시험 2차시험 발전방안’에서는 국내 26개 전문과목 학회를 대상으로 현행 전문의자격 2차시험에 대한 각 학회별 인식도 및 2차 시험 개선방안에 대한 설문조사를 실시하였다. 54회 전문의 자격시험에서 각 전문학회는 과목별 특성에 따라, 슬라이드 시험, 구술시험, 병리판독, 방사선사진판독, 동영상 시험, 종합적 환자취급, OSCE, CPX 등 다양한 유형의 2차시험을 시행하고 있었다(Table 4). 현재 시행되는 2차시험 유형은 단순 슬라이드시험이 대부분이었으며, 이는 지식수준 혹은 인지영역 평가에 그쳐 실기시험의 취지에 부합하지 못하였다. 일부 전문학회에서 시행된 구술시험의 경우, 문제 당 배정된 시험시간이 수 분 이내로 외국의 구술시험에 비해 시험시간이 짧고, 평가점수의 객관적 기준이 없어서, 신뢰도를 충분히 확보하지 못하는 것으로 평가되었다.

설문조사 결과에서는, 2차시험의 필요성에 대한 질문에 24개의 학회가 필요하다고 응답한 반면, “현재 2차시험이 시험방법 상, 실기시험이라고 판단하십니까?”라는 질문에는 8개 학회만이 실기시험이라고 응답하였으며, 이 경우 주로 구술시험 혹은 OSCE, CPX의 유형이 해당되었다. 또한, 13개 학회가 실기시험과 비슷하다고 답하였고, 4개 학회는 실제 실기시험이 아니라고 응답하였다. 대부분의 전문학회가 현재 시행되는 2차 시험의 문제점을 인식하고, 개선의 필요성을 느끼고 있으나, 현실적으로 개선이 어렵다는 학회가 14개 학회나 되었다. OSCE, CPX 형태로의 전환을 요구 받을 경우, 개선 가능하다는 학회가 7개였지만, 불가능하다는 학회가 14개나 되었다. 불가능하다고 답한 14개 학회의 경우, 인력 및 예산의 한계, 전문과의 진료특성 상, 표준화 환자 확보의 어려움, 전문과의 특성 상, 자료판독 이외의 실

Table 4. The step 2 skill examination methods in specialty certifying examination, Korea (2013)

Examination methods	
Picture test using slide (19)	Family medicine, internal medicine, anesthesiology, urology, obstetrics and gynecology, pediatrics, psychiatry, neurology, ophthalmology, radiology, general surgery, emergency medicine, orthopedic surgery, rehabilitation medicine, dermatology, occupational and environmental medicine, thoracic surgery, laboratory medicine, nuclear medicine
Oral examination (9)	Pathology, neurology, neurosurgery, psychiatry, dermatology, orthopedic surgery, thoracic surgery, otorhinolaryngology, preventive medicine
Patient management program (7)	Internal medicine, orthopedic surgery, neurosurgery, neurology, otorhinolaryngology, psychiatry, radiation oncology
Reading of pathology slides or X-ray (10)	Internal medicine, pathology, radiology, general surgery, dermatology, orthopedic surgery, neurosurgery, rehabilitation medicine, pediatrics, nuclear medicine
OSCE (3)	Plastic surgery, neurology, otorhinolaryngology
CPX (4)	Family medicine, plastic surgery, ophthalmology, otorhinolaryngology

OSCE, objective structured clinical examination; CPX, clinical practice examination.

Table 5. The self-awareness and the developmental plan on the step 2 skill examinations in specialty certifying examination, Korea

Survey questions	Korean boards of medical specialties (n)
Degree of satisfaction with the current step 2 examination methods	
Very satisfied	1
Satisfied	12
Neutral	4
Dissatisfied	0
Degree of necessity of the step 2 examination in board certification	
Very necessary	8
Necessary	16
Neutral	1
Not necessary	0
The appropriateness of the current step 2 examination methods to assess clinical skill and competence	
Agree	8
Neutral	13
Disagree	4
The necessity on improving of the current step 2 skill examination methods	
Necessary, and we have a plan to improve it	8
Necessary, but it is difficult	14
Not necessary	1
The possibility that convert the skill examination methods to OSCE/CPX	
Possible	7
Impossible	14
We have already performed OSCE/CPX	4
The necessity of combining the work based assessment with the skill examination	
Necessary	9
Neutral	4
Not necessary	12

OSCE, objective structured clinical examination; CPX, clinical practice examination.

기시험이 불필요하다는 것을 이유로 제시하였다. 2차시험과 수험평가의 연계 필요성에 대해서는 9개 학회가 동의하였다(Table 5).

각 전문학회의 향후 2차시험 개선계획으로는, OSCE, CPX 형태로의 전환, 구술시험의 보완, 환자, 수술 동영상자료 개발, 수련기간 중 기본술기 이수 인증 등의 여러 의견이 있었으나, 비용, 장소, 표준화 환자 확보 및 기술적 문제 등 계획을 현실화하는 것에 어려움을 느끼는 경우가 많았다. 따라서, 국내 전문의 자격 2차시험이 임상기술과 태도에 대한 평가를 강화할 수 있도록 내실화하기 위해서는, 각 전문학회의 특성을 고려하여, 적용가능한 개선방안을 수립하고, 이를 실현하기 위한 실질적 지원이 필요할 것으로 판단된다.

국내 전문의자격시험 실기시험 발전방향의 원칙과 발전방안

1. 실기시험 발전방향의 원칙

대한의사협회 연구용역 보고서 ‘전문 의자격시험 실기시험 개선방안’에서는 의학교육평가로서의 전문의자격시험의 의의를 살리기 위해 실기시험에 대한 발전방안 원칙을 다음과 같이 제시하였다.

첫째, 전문의자격시험 실기시험은 의학교육평가 원칙에 합당하여야 한다. 즉 타당도, 신뢰도, 객관성 및 공정성을 강화가 전제되어야 한다.

둘째, 전문의자격시험 중 실기시험은 전문의로서 전문진료능력(professional

competence)의 평가에 주안점을 두되, 1차시험인 필기시험과 다른 영역인 실기시험에 합당하여야 한다. 즉 교육목적에 따라 기술과 태도 또는 행동영역을 평가하는 시험이어야 한다.

셋째, 교육평가원칙에 맞추어 실기시험을 시행할 경우 많은 인력과 시간 그리고 비용이 필요하므로, 전문학회별 전문적 특성 및 준비여건을 고려하여 시험형식과 실시시기 등에 대해 전문학회의 자율성과 다양성을 인정하여야 한다.

넷째, 전문의자격시험 실기시험의 준거 설정(합격선)은 다양한 시험형식에 따른 적절한 방법을 인정하여야 한다. 전문의고시는 준거평가인데 현재 전문의고시의 준거점수는 '전문회의 수련 및 자격 인정 등에 관한 규정 시행규칙' 제13조 2항에 의하면 총점의 60% 이상으로 규정되어 있다. 하지만 실기시험은 형식에 따라 적합한 준거설정 방법(합격선)이 각각 달라질 수밖에 없으므로 학회의 전문적 자율성을 인정하는 것이 필요하다. 하지만 시험의 객관성과 공정성의 원칙을 훼손하지 않는 범위에서 자율성이 인정되어야 한다.

2. 국내 전문의자격시험 실기시험의 단계적 발전방안

'전문회의 수련 및 자격 인정 등에 관한 규정 시행규칙' 제13조 2항에 의하면 '전문의자격시험은 1차시험과 2차시험으로 구분하되 1차시험은 필기시험으로 하고 2차시험은 실기시험 또는 구술시험으로 한다. 1차에 합격한 사람만이 2차시험을 볼 수 있다'라고 규정되어 있다. 현행 법 규정하에서는 실기시험을 전공의 수련과정의 평가인 WBA로 대체한다든지, 학회 특성상 실기시험이 필요 없는 경우에도 필기시험 단독으로 시행할 수 없다. 따라서 실기시험의 전문성, 다양성, 자율성에 근거한 발전 방향의 개선을 위해서는 이 법규정이 개정될 필요가 있다. 대한의사협회 연구용역 보고서 '전문의자격시험 실기시험 개선방안'에서는 단계적 개선방향을 법 개정 전 단기적 발전 방안과 법개정을 포함한 장기적 발전방안으로 나누어 제시하였다.

1) 법규정 개정 전 단기적 개선방안

첫째, 슬라이드시험은 사진 및 그림을 인쇄물로 제시하는 필기시험으로 통합한다. 국내 전문의 실기시험에 흔히 사용하는 슬라이드시험이나 빔프로젝트시험은 사진 및 그림을

제시하는 방식이 슬라이드나 빔프로젝트일뿐 진정한 의미의 실기시험이 아니며, 지식수준 또는 인지영역 평가시험이다. 즉 객관식 문항의 필기시험과 동일한 시험이므로 두 가지 형식의 시험으로 따로 시행할 타당성이 부족하다. 사진 및 그림을 인쇄물 형식으로 제시하면 슬라이드로 제시하는 것보다 수험자가 해석하고 판정할 수 있는 시간이 충분하여 오히려 신뢰도가 높아진다. 또한 슬라이드로 제시하는 시험문항 수는 한정될 수 밖에 없고, 적은 문항 수로 인한 신뢰도 저하는 전문의자격시험에서 합격률 변동의 주요 원인 중 하나였다. 따라서 사진 제시형 필기시험으로 통합하면 이 한계를 극복할 수 있다.

둘째, 구술시험의 신뢰도 제고 방안을 고려한다. 현재 국내 전문의 흔히 사용하는 구술시험은 태도를 평가할 수 있으나 주된 평가영역이 인지영역평가이고, 술기를 평가하는 실기시험으로 아니므로, 교육평가원칙에 준하여 장기적으로 지양해야 하는 것이 적절하다. 여건 상 시행한다고 하더라도 구조화된 구술시험으로 신뢰도를 높이는 방식으로 보완이 필요하다. 또한 구술시험이라 하더라도 인지영역 평가보다 행동영역을 평가하는 방식, 즉, 전공의 과정의 증례 수집, 수집 후 증례에 대해 토론하는 구술시험으로 개선할 것을 검토해 봐야 한다.

셋째, 실질적인 실기시험 확대를 위해 여건이 준비된 학회는 객관성과 신뢰도가 검증된 실기시험형식인 OSCE/CPX을 도입을 노력이 필요하고, 제도적인 장려정책이 필요하다.

2) 법개정 후 장기적 발전방안

실기시험 본연의 형식으로 전환할 경우 학회의 전문적 특성과 준비 여건에 따라 실시시기, 내용, 형식 등에 대해 전문학회의 다양성을 반영하여 선택할 수 있어야 한다. 현행 방식 이외에도 전문학회가 실기시험을 WBA에 기초한 전공의 수련 형성평가로 전환 혹은 반영하거나, 수련 후 질 관리를 반영한 PBA로 변화시키거나, 혹은 필요 없다면 아예 필기시험만 시행하는 방식 중에 선택의 자율성을 확대하는 것이다.

첫째, 전공의 수련평가 반영으로 전환을 허용해야 한다. 전문의자격시험 실기시험을 OSCE, CPX, Mini CEX, WBA 등 다양한 형식의 실질적인 실기시험으로 전환할 경우, 현행 시험기간인 1월 초에 전부 시행하기에는 인력, 시간 등의 현

실적인 한계가 있다. 또한 전문의자격시험은 불합격으로 인한 불이익이 매우 큰 고부담시험이다. 현재 전공의 수료예정자는 1회 시험에 의한 탈락위험을 줄이기 위해 전공의 수련과정 중 마지막 수개월을 시험 준비에만 몰두하여 수련을 받지 않고 있는 실정이다. 이러한 현실을 타개하기 위해서는 전문의자격시험 중 실기시험은 행동영역 평가 중 WBA로 전환하여 수련과의 유기적인 관계를 높이는 방안에 대해 고려해야 한다.

둘째, 법적으로 전문의자격시험으로 필기시험 단독으로 시행할 수 있도록 허용해야 한다. 현행 슬라이드(빔프로젝트)시험을 폐지하고 사진 인쇄물 제시형 필기시험으로 전환한다면, 전문학회 특성상 신뢰도와 타당도를 갖춘 실질적인 실기시험을 시행하기 어렵거나, 아예 시행 필요가 없는 학회도 있다면 필기시험 단독 시행할 수 있게 하는 것이다.

셋째, 실기시험은 전문의 면허 획득 후 질 관리 평가방식으로 전환을 허용해야 한다. 미국의 전문의시험에서 흔히 사용하는 practice base oral examination(필기시험 합격 후 전문의 진료사례 수집 후 토론 평가)같은 형식을 참조하여 전문의 진료 질 평가 방식 활용하는 것이다. 예를 들면 필기시험이 끝난 후 전문의 1차면허를 부여하고, 전문의 자격으로 활동 후 일정기간 지난 후 증례리스트, 포트폴리오를 제출하여 진료사례의 질 관리 평가로 활용하는 것이다.

결론

국내 전문의자격시험은 의학교육에 있어서 최상위 시험으로, 1, 2차시험으로 나누어 시행된다. 1차시험은 필기시험, 2차시험은 실기시험으로 구성되나, 현행 국내 26개 전문학회 2차시험의 대부분은 사진 혹은 그림을 제시하는 슬라이드 시험의 형태이다. 이는 실질적인 임상능력을 평가하는 실기시험이라기보다는 인지영역을 평가하는 필기시험의 연장이며, 일부에서 시행되는 구술시험은 짧은 시험시간과 객관적 채점기준의 부재로 신뢰도가 떨어지는 경우가 많았다. 따라서, 국내 전문의자격 2차시험이 임상 수행 역량 평가라는 실기시험 본연의 목적에 부합하는 방향으로 발전하기 위해서는

단계적 개선이 필요하며, 대부분의 전문학회에서도 개선의 필요성에 대해 동의하고 있다. 우선 현행 구술시험의 신뢰도를 높이는 노력이 필요하며, 객관성과 신뢰도가 검증된 실기시험형식인 OSCE/CPX을 도입하는 것을 고려해야 한다. 장기적으로 전문학회의 전문성, 자율성이 반영되고 1회시험으로 합격이 결정되는 부담을 낮추고, 시험과 수련내용의 유기적인 관련성을 높이는 실기시험으로 발전하기 위해 WBA인 전공의 수련평가를 반영하거나 대체하는 방안을 강구하기 위해서는 우선 관련 법 개정이 선행되어야 한다.

찾아보기말: 전문의자격시험; 수행평가; 실기시험; 직무중심평가

ORCID

Jung Jin Cho, <http://orcid.org/0000-0001-8871-8858>

Hye Mi Noh, <http://orcid.org/0000-0003-2112-1613>

Seung Ho Kim, <http://orcid.org/0000-0002-2733-9763>

Ho Kwon, <http://orcid.org/0000-0001-7471-0804>

Young Min Park, <http://orcid.org/0000-0002-3631-0807>

Byung Min Choi, <http://orcid.org/0000-0002-0931-2353>

REFERENCES

1. Crossley J, Humphris G, Jolly B. Assessing health professionals. *Med Educ* 2002;36:800-804.
2. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 Suppl):S63-S67.
3. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001;357:945-949.
4. Amin Z, Khoo HE. Basics in medical education. London: World Scientific; 2003.
5. Driessen E, van Tartwijk J, van der Vleuten C, Wass V. Portfolios in medical education: why do they meet with mixed success? A systematic review. *Med Educ* 2007;41:1224-1233.
6. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach* 2007;29:855-871.
7. Postgraduate Medical Education and Training Board. Developing and maintaining an assessment system: a PMETB guide to good practice [Internet]. London: The Assessment Working Group; 2007 [cited 2014 Mar 14]. Available from: http://www.gmc-uk.org/Assessment_good_practice_v0207.pdf_31385949.pdf.

8. Van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ* 2005;39: 309-317.
9. Evgeniou E, Peter L, Tsironi M, Iyer S. Assessment methods in surgical training in the United Kingdom. *J Educ Eval Health Prof* 2013;10:2.
10. Mitchell C, Bhat S, Herbert A, Baker P. Workplace-based assessments of junior doctors: do scores predict training difficulties? *Med Educ* 2011;45:1190-1198.
11. Brennan TA, Horwitz RI, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification status in the quality movement. *JAMA* 2004;292:1038-1043.
12. The American Board of Medical Specialties. About board certification [Internet]. Chicago: The American Board of Medical Specialties; 2013 [cited 2014 Mar 14]. Available from: http://www.abms.org/About_Board_Certification/.
13. The American Board of Physician Specialties. ABPS Boards of Certifications [Internet]. Tampa: The American Board of Physician Specialties; 2014 [cited 2014 Mar 14]. Available from: <http://www.abpsus.org/abps-medical-board-certifications>.
14. The College of Family Physicians of Canada. Certification examination in family medicine [Internet]. Mississauga: The College of Family Physicians of Canada; 2014 [cited 2014 Mar 14]. Available from: <http://www.cfpc.ca/FMExam/>.
15. The Royal College of Physicians and Surgeons of Canada. Examination guidelines for registered candidates [Internet]. Ottawa: The Royal College of Physicians and Surgeons of Canada; 2013 [cited 2014 Mar 14]. Available from: http://www.royalcollege.ca/portal/page/portal/rc/credentials/start/exams/candidate_information.
16. The Royal College of General Practitioners. MRCGP exam overview [Internet]. London: The Royal College of General Practitioners; 2014 [cited 2014 Mar 14]. Available from: <http://www.rcgp.org.uk/gp-training-and-exams/mrcgp-exam-overview.aspx>.