

제1 회 정형외과 전공의 평가시험 결과의 분석

대한정형외과학회

석 세 일 · 장 의 열 · 유 명 철

서울대학교 의과대학 정형외과학교실

정 문 상

— Abstract —

The First In-Training Examination for Residents in Orthopaedic Surgery

Se Il Suk, M.D., Ik Yull Chang, M.D., Myung Chul Yoo, M.D. and Moon Sang Chung, M.D.

The Korean Orthopaedic Association

On July 5, 1980, the first in-training examination for Korean orthopaedic residents was performed to give each of them an evaluation of his general level of orthopaedic knowledge, and to provide their program director with a comparative basis for evaluation of their programs in various subject areas. Forty one of the 42 residency programs in this country presented in the examination. The results of the examination and the survey of public opinion performed after the examination are as follows.

1. Two hundreds and ninety one of the 312 Korean orthopaedic residents at all level of training participated in the examination, so that the participation rate was 93.3 percent. According to the survey, the respondents wanted to continue the in-training examination in 77 percent.
2. The examination was not sufficient in its reliability because the questions were small in number, and it was deficient in its objectivity since there were 40 percent of the subjective questions. The test was not enough in its practicability because the questions of applying the knowledge clinically such as problem solving one were relatively insufficient in its number.
3. The mean scores of the residents are increasing regularly with increasing years, which can be interpreted as the questions are somewhat adequate for the test and the orthopaedic training in Korea is efficient one. But the standard deviations of each mean is large, which means the residents level of the orthopaedic knowledge and his training are greatly varied. The scores of the senior residents are not yet reach a plateau and the scores of the junior residents are inferior to that of the 1966 American counterpart.
4. According to the survey, the majority wants to open the questions with their referencies after the examination, but wants to neither open his score nor reflect his score in that of the Korean orthopaedic board examination.

Key words : In-Training Examination.

I. 서 론

대한정형외과학회에 가입된 각 년차 전공의 291 명
에 대한 제 1 회 정형외과 전공의 평가시험이 1980 년

7월 5일 시행되었다. 본 시험은 (1) 한국 정형외과
전공의들의 지식과 기술에 관한 능력을 평가하고, (2)
본 시험을 통하여 전공의들이 공부하도록 자극하며, 더
나아가서는 이번 시험에서 노출된 그의 학문에의 약점
을 보강하게하여 전공의의 실력을 배양하며, (3) 각 수

련병원의 정형외과 과장들이 소속 전공의들의 실력을 인지하므로써 그들의 교육제도의 단점과 약점을 보강하게 하고, 또한 본 시험이 계속될 경우에 대한정형외과 학회는 한국 정형외과 교육의 목표와 교육 계획등을 제고하고 개선할 수 있는 좋은 자료로 사용될 수 있게 하는 것이다(4,5,7,8,9).

전공의 평가시험은 1963년 American Board of Orthopaedic Surgery에서 처음으로 실시하였고, 1964년 American Board of Neurological Surgery에서 두번째로 시행하였으며 그후 안과, 피부과, 산부인과 등 미국의 각 과에서 널리 시행하게 되었다(3,5,7,8,9,11,12). 우리나라에서는 1977년 대한방사선과학회에서 전공의 평가시험을 처음으로 시행하였으며, 대한정형외과 학회에서는 1980년 7월 제 1회 정형외과 전공의 평가시험을 시행하였다. 대한정형외과 학회에 소속된 42개 정형외과 전공의 교육기관중 41개의 수련병원이 적극적으로 참여한 가운데 본 시험은 시행되었으며, 각 전공의의 점수와 년차별 그리고 분야별 평균은 해당자의 소속 과장을 경유하여 비공개리에 시험자에게 전달되도록 배려하였다. 본 논문은 전공의 평가시험의 장단점, 효과등을 고려하여 앞으로 시험의 지속 여부를 논의하고, 시험의 출제, 채점, 관리등을 검토하기 위하여 고안되었다.

II. 시험방법

본 시험은 전공의들의 교육을 도울기 위하여 시행된 것이므로 강제성이 없이 자율적으로 전공의들이 참석하게 하였고, 시험의 결과에 의한 포상이나 처벌은 일체 고려되지 않았으며, 대한정형외과 전문의 시험과는 아무런 연관이 없이 시행되었다.

학회 본부와 4개 지회에서 본 시험은 동시에 시행되었으며, 수련의 방법이나 수준이 다른 모든 수련병원의 모든 년차의 정형외과 전공의를 대상으로하여 설정된 총응시 예정자 312명중 291명이 시험에 참석하여 93.3%의 응시율을 나타내었다.

출제는 교직자 40여명이 분야별로 출제하게 하고 6명의 선택위원이 문제를 선택하였으며, 8명의 교직자가 분야별로 분담하여 채점하였다.

표 1. 제 1회 정형외과 평가시험

시 기	: 1980. 7. 5
응시에정자	: 312명
결 석 자	: 21명
응 시 자	: 291명
응 시 율	: 93.3%

문제는 일반, 성인, 소아정형외과학, 외상, 수부, 정형외과 해부학, 병리학, 생리 및 생화학 그리고 생역학과 재활의 9개 분야로 나누어서 출제하였다. 분야별 출제율은 일반정형외과학과 외상이 각각 20%, 성인정형외과학이 15%, 소아정형외과학, 정형외과 해부학, 병리학이 각각 10%, 그리고 수부, 생리 및 생화학, 생역학 및 재활이 각각 5%를 차지하였다.

표 2. 정형외과 분야별 문제 출제율

분	야	문 제 수	점 수 (%)
일	반	20	24(20)
성	인	15	18(15)
소	아	10	12(10)
외	상	20	24(20)
수	부	5	6(5)
해	부	10	12(10)
병	리	10	12(10)
생 리 및 생 화 학		5	6(5)
생역학 및 재활의학		5	6(5)
합	계	100	120(100)

문제의 형식은 서술형의 주관식이 20문제에 40점, 선다형 객관식이 80문제에 80점으로 합계 100문제에 120점으로 채점하고 이를 100점으로 환산하여 개인별 평균점수로 하였다.

III. 시험의 결과

시험의 결과를 분석하기 위하여 첫째로 9개 분야별로 개인별 성적을 채점하였으며, 여기에서 각 년차 및 수험자 전체의 분야별 평균을 산출하였다. 또한 개인의 평균에서 년차별 평균과 표준편차를 산출하였고, 수험자 전체의 평균을 구하였다. 둘째로 시험후에 여론조사를 실시하여 이를 분석하였다.

수험자 전체의 평균은 52점이었으며, 년차별 평균과 표준편차는 1년차 88명에 46 ± 8.6 , 2년차 71명에 48 ± 9.6 , 3년차 64명에 56 ± 10.3 , 그리고 4년차는 67명에 65 ± 9.4 점이었다. 그러므로 년차별로 8~9점(평균 8.3점)의 차이를 나타내어 정형외과의 수련과정이 효과적으로 진행되고 있음을 보여주었다. 분야별로 가장 우수한 분야는 평균 64점인 외상이었으며, 소아정형외과학과 생역학 및 재활이 각각 40점 및 41점으로 가장 열등하였다.

시험후 시행한 여론조사에는 총 291명중 82명이 응답하였다. 이들 중 전공의 시험이 많은 도움이 되었다고 답한 사람이 24명에 29%, 어느 정도 도움이 되었

표 3. 제 1 회 정형외과 전공의 평가시험 결과표

분야 년차	일반	성인	소아	의상	수부	해부	병리	생리 생화학	생역학 재활	최하	최고	평균	표준 편차
1년차	36	40	27	51	31	44	43	41	35	19	64	40	8.6
2년차	42	49	36	61	42	49	51	44	39	25	73	48	9.6
3년차	51	58	47	70	49	61	56	48	46	23	79	56	10.3
4년차	59	67	54	80	66	67	66	54	45	36	84	65	9.4
평균	46	52	40	64	46	54	53	46	41	19	84	52	—

* 분야별 성적은 100점 만점으로 환산된 것이므로 분야별 합산과 평균치와는 일치하지 않음.

다는 전공의가 41명에 50%, 그리고 도움이 안되었다는 답이 14명에 17%이었으며, 총 응답자의 77%는 전공의 시험이 계속되기를 원하고 있었다.

시험문제의 여러가지 요소에 대한 응답에서는 금후의 시험시 문제수는 100문제 이내가 좋겠다고 하는 주장이 42%, 100문제 이상이어야겠다는 의견이 58%이었으며, 대부분의 응시자는 이번 시험이 공정하였다고 그 난해도는 무난하였으며 분야별로 적당히 분포되어 있었다고 답하고 있었다. 응시자의 56%는 금후 객관식 문제가 증가하기를 희망하였고, 44%는 주관식 문제가 늘기를 희망하였다.

표 4. 전공의 시험의 유효성

	유효 인원	성 (%)
많은 도움	24	(29)
약간 도움	41	(50)
도움 안됨	14	(17)
모르겠다	3	(4)
합 계	82	(100)

본 전공의 평가시험의 분야별 채점에서 소아정형외과학과 생역학 및 재활은 그 성적이 매우 나빴음은 이미 상술했거니와, 여론조사에서도 생역학 및 재활, 생리 및 생화학, 병리학 그리고 소아정형외과학의 순서로 어려워했다고 응답하고 있었다.

전공의 시험 결과의 처리에 관해서는 응답자의 93%가 문제를 참고문헌과 함께 공개하기를 원하고 있었으며 90%는 성적이 공개되기를 원하지 않았고, 62%는 전공의 시험이 전문의 시험에 반영될 수 없다고 단호하게 주장하였다.

IV. 고 안

제 1 회 정형외과 전공의 평가시험은 전공의들의 높은

응시율과 더불어 대부분의 전공의들이 본 시험이 유의했다고 생각하는 점, 그리고 각 전공의의 실력을 평가하여 그들에게 공부하도록 자극을 주고 약점을 분석해 주었으며, 그들이 소속된 교육기관의 교육 목표 내지는 교육 계획을 반성하고 재고할 수 있는 기회를 부여하여 큰 성과를 얻었다고 할 수 있겠다.

그러나 모든 시험은 연속적인 과정으로서 하나의 시험을 시행한 다음에는 그 시험의 목적과 평가 기준을 재고하고, 현재까지 시행되어온 교육 방법을 반성하게 하며, 나아가서는 보다 나은 평가 방법을 강구하게 하고 전공의 교육의 새로운 목표와 계획을 수립하게 한다는 점에서 본 전공의 평가시험은 더욱더 그 의의를 찾을 수 있을 것이다²⁾. 이러한 면에서 고려한다면 제 1 회 정형외과 전공의 평가시험은 해결이라기 보다는 오히려 많은 문제점을 제시하고 있다.

무릇 모든 시험문제는 (1)신빙성(Reliability), (2)객관성(Objectivity), (3)타당성(Validity) 그리고 (4)실용성(Practicability)의 네가지 요소를 갖추어야 한다고 한다²⁾. 시험자의 실력을 보다 정확하게 판정하기 위하여는 즉 통계적인 신빙도를 높이기 위하여는 문제의 수를 늘려야 한다. 이러한 관점에서 100문제는 그 수가 너무 부족하며 또한 정형외과의 전 영역을 완전히 포괄하기 힘들다. 본 시험에서는 주관식의 점수가 40점에 이르고 있어 막중한 비중을 차지하고 있다. 따라서 어떤 질량을 자못 이해하고 있는 전공의가 단지 암기하지 못하였거나 문장의 구성 능력이 부족하여 낮은 점수를 맞았을 가능성이 농후하다. 이는 문제의 객관성이 부족하여 파생되는 부작용으로서 금후에는 주관식의 비율을 줄여 객관성을 높여야 할 것이다. 또한 출제된 문제의 타당성에 대하여도 시험후에 시행한 여론조사에서 상당수의 응답자는 몇몇 문제에 타당성을 의심하고 있었으며, 그들 문제에 대한 검토에서 타당성이 어느 정도 결여되어 있음을 발견할 수 있었다.

선택된 문제들의 내용은 전부 전공의의 지적 분야

(Cognitive Domain)를 취급한 것으로서, 이중 지식(Knowledge)에 관한 문제가 44%, 이해력(Understanding)이 36% 그리고 응용력(Application)에 관한 문제가 20%이었다. 응용력에 관한 문제는 분석 및 해석(Analysis and Interpretation)에 관한 것이 40% 그리고 문제해결 능력(Ability to solve problem)을 검사한 것이 60%이었다. 또한 병력 채취, 자료의 수집과 정리, 병특지 기재 능력이나 수술의 방법이나 과정 그리고 소독법이나 기구의 사용법 등 정형외과 의사로서 꼭 필요한 실용적인 문제가 출제되지 않았다. 더구나 정서적 분야나 책임감등 비지적 분야(Noncognitive Domain)도 출제에서 제외되어 있었다^{1,6,10)}. 따라서 본 시험은 전공의의 지식

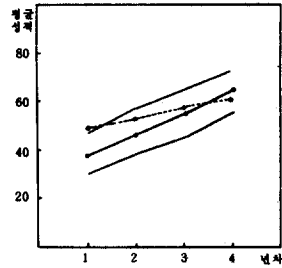
표 5. 선택된 문제의 내용

분	야	문제수(%)
1. 지적분야(Cognitive Domain)		
ㄱ.	지식(Knowledge)	44
ㄴ.	이해력(Understanding)	36
ㄷ.	응용력(Application)	20
1)	분석 및 해석 (Analysis & Interpretation)	8
2)	문제 해결 능력 (Ability to Solve Problem)	12
3)	기타(병력, 자료정리, 기재, 수기등)	0
2. 비지적분야(Noncognitive Domain)		
		0

과 이해력만을 검사한 비실용적인 시험이었던 인상을 짙게 하고 있다. 그러므로 금후의 전공의 시험에서는 분야별 출제에 그치지 말고 문제의 내용을 포함한 문제의 제반 요소들을 중요시하여 타당성과 객관성이 높은 문제를 선택하고, 전공의의 단순한 지식이나 이해력 검사를 지양하고 실제로 환자를 처리하는데 필요한 실용성이 풍부한 문제 즉 문제를 위한 문제가 아닌 문제를 다수 출제함과 더불어, 문제의 수를 늘여 실험도를 높임으로써 전공의의 정형외과학 전반에 관한 능력을 보다 정확하게 판정할 수 있을 것으로 사료된다.

본 시험의 년차별 평균점수는 년차가 증가함에 따라 8~9점씩 규칙적으로 증가하고있어 출제가 비교적 적절하였으며, 정형외과 전공과정이 효과적으로 시행되고 있음을 나타내었다. 그러나 년차별 평균의 표준편차는 매우 커서 10 점대에 육박하고 있었으므로, 우리나라의 정형외과 전공의는 그 실력의 차이가 상당히 심하고 수련의 수준과 정도도 크게 다르다는 사실을 암시하고 있었다. 그러므로 전공의의 실력 수준을 균일하게 평준화

표 6. 년차별 평균성적의 비교



- 한국, 제 1회 전공의 평가시험 평균
- 한국, 제 1회 전공의 평가시험 표준편차
- 미국, 정형외과 전공의 평가시험 평균, 1966

하는 일은 우리나라 정형외과 교육의 큰 문제일 것으로 사료된다.

년차별 평균과 표준편차를 도표로 표시하면 표 6과 같다. 이 년차별 평균을 1966 년도에 시행한 미국 정형외과 전공의 평가시험과 비교하여보면 우리나라의 곡선은 미국의 그것과는 달리 아직 4년차에서 고원현상(Plateau)에 도달하지 못하였다. 전공의의 실력이 고원현상을 보이기 시작하는 시점에서 수련을 끝내는 것이 적당할 것이라는 전지에서 생각해 보면 미국의 정형외과는 4년의 정형외과 전공의 교육기간으로서 적당하다고 사료되나 우리나라에서는 정형외과 전공의가 4년으로써 적당한 실력에 도달하였다고 말할 수 없다. 또한 이들 곡선에서 4년차의 평균성적은 우리나라와 미국이 비슷하나, 그 차이는 저년차로 갈수록 심해져서 1년차에서는 한국의 평균이 미국에 비해 한국의 표준편차 보다도 더 열등감을 보여주고 있다. 저년차의 점수가 미국에 비해 뒤진다는 사실은 세가지 측면에서 해석될 수 있겠는데, 첫째는 양시험의 문제의 차이에서 발생하였을 가능성이 있고, 둘째는 한국의 전공의 시험이 시기적으로 너무 일러서 발생하였을 가능성이 있으며, 셋째로 우리나라 저년차 전공의의 실력이 뒤떨어져서 발생하였을 가능성도 있다. 현재로는 상기 3자 중 어느 것이 주요 요인인지 판정할 수 없으나 만약 이러한 현상이 우리나라 저년차 전공의의 뒤떨어진 실력에 기인하는 것이라면 이는 정형외과 학생 교육을 포함한 기초적인 교육이 잘못되었다는 것을 시사하는 소견으로서 이는 금후 연구를 거듭하여 해결하여야 할 전공의 교육 및 시험제도의 과제일 것으로 사료된다.

본 정형외과 전공의 평가시험의 분야별 채점에서 소아정형외과학과 생역학 및 재활은 그 성적이 매우 나빴음은 상술한 바와 같다. 여론조사에서도 생역학 및 재활, 생리 및 생화학, 병리학 그리고 소아정형외과학이 어려운 것으로 나타나 있는 바, 이러한 분야에 대한 전

공의 실력을 높이기 위한 특별한 배려가 필요할 것이다.

상술한 바와 같은 전공의 실력 수준의 문제점 즉 실력의 평준화나 일반적으로 실력이 뒤떨어지는 분야에 대한 보완책으로써 교환 전공의 제도, 교환 교수 제도, 또는 현재의 수련병원 단위에서의 수련이 아니고, 대학병원을 중심으로한 교육그룹의 설정등 여러가지 제도를 고려해 볼 수 있으나, 이는 수차의 시험을 치른후 점진적으로 생각할 문제이며, 현 단계에서 논의될 문제는 아닌 것으로 사료된다.

전공의 평가시험이 계속되기를 희망한 여론조사의 응답자는 전체의 77%이었고, 반대자는 23%이었는데, 이들이 평가시험을 반대한 이유들로는 시험 준비 때문에 환자의 진료 및 기타 업무 수행의 기피현상, 년차별 유대관계의 파괴, 정서적과 시간적인 부담감, 시험을 포기하여 공부를 더 안하게 됨, 그리고 시험 위주의 교육 가능성이 농후함등을 지적하고 있었다. 또한 금후 문제의 개선점에 관하여는 년차별로 다른 문제를 출제하여 년차별 유대관계를 원활하게 할 것, 실질적이고 임상적인 문제, 객관성이 풍부한 문제, 문제 해결형이나 이해를 요구하는 문제가 늘 것을 바라고있었으며, 골관절의 결핵이나 화농성 골수염등 흔한 질환에서 더 많이 출제되기를 바라고 있었다.

전공의 시험 결과의 처리에 관하여 응답자의 대부분이 문제의 공개를 원하고 있었으므로 금후에는 시험을 치른후 문제지는 전공의에게 지급하는 것이 타당할 것으로 사료되며, 성적의 공개나 전문의 시험에의 반영은 반대한 응답자가 많았고 또 부작용이 많고 심각할 것으로 사료되므로 고려치 않는 것이 바람직할 것이다.

V. 결 론

대한정형외과학회에서는 1980년 7월 제 1회 정형외과 전공의 평가시험을 시행하여 다음과 같은 결론에 도달하였다.

1. 총응시 예정자 312명중 291명이 시험에 응하여 93.3%의 응시율을 나타내었으며, 시험후 시행한 여론조사에서는 77%의 응답자가 전공의 평가시험이 지속되기를 희망하고 있었다.

2. 분야별로 출제된 문제들은 그 문제수가 적어서 신뢰도가 충분하지 못하였으며, 주관식의 배점이 너무 많아 객관성이 부족하였고, 또한 문제의 내용에 있어서 지식의 응용력을 검사하는 문제가 비교적 적어서 실용도가 크다고 말할 수 없었다.

3. 본 시험의 평균성적은 년차별로 규칙적으로 증가하고있어 출제가 비교적 적절하였다는 것과 정형외

과 전공의 교육이 효율적임을 나타내었다. 그러나 년차별 평균의 표준편차는 매우 커서 전공의의 교육수준이 여러가지 단계에 있었음을 알 수 있었다. 4년차의 평균성적은 아직 교원현상에 도달하지 못하고 있었으며, 저년차의 성적은 미국의 그것에 비해 낮았다.

4. 시험후 시행된 여론조사에 의하면 대부분의 응답자는 시험문제를 참고문헌과 함께 공개하기를 원하였으나, 성적의 공개나 전문의 시험에의 반영에는 난색을 표시하고 있었다.

REFERENCES

1. Charvat, J., McQuire, C. and Parson, V. : *A Review of the Nature and Use of Examinations In Medical Education*. World Health Organization, Geneva, 1968.
2. Gault, N. L. : 평가방법 - 제 1회 한국의학교육 세미나 - 의학교육 : 122~140, 1971.
3. Gregory, C.F. : *Orthopaedics and the Impact of Learning Theory*. J. Med. Educ., 44: 777-783, 1969.
4. Gregory, C.F. : *Some Observations on the Postgraduate Medical Education of Orthopaedic Surgeons*. J. Bone Joint Surg., 56-B: 573-579, 1974.
5. Hubbard, J.P., Furlow, L.T. and Matson, D.D. : *An In-Training Examination for Residents as a Guide to Learning*. New Engl. J. Med., 276-8: 448-451, 1967.
6. Keck, J.W. and Arnold, L. : *Efficiency of Cognitive, Noncognitive Measures in Predicting Resident-Physician Performance*. J. Med. Educ., 54: 759-765, 1979.
7. Matson, D.D. : *An In-Training Evaluation of Residency Training Programs and Trainees*. J. Med. Educ., 41: 47-51, 1966.
8. Meskaukas, J.A., Newton, M. and Russell, K.P. : *The 1973 In Training Examination in Obstetrics and Gynecology*. Obstet. Gynecol., 44-3: 463-468, 1974.
9. Miller, G.E. : *The Orthopaedic Training Study*. JAMA, 206-3: 601-606, 1968.
10. Rice, R. G. : 한국에서 사용할 수 있는 몇가지 평가방법 - 제 1회 한국의학교육세미나 - 의학교육 : 141~149, 1971.
11. Rubin, M.L. : *The Ophthalmology Resident In-Training Examination*. Amer. J. Ophthal., 67-1: 70-79, 1969.
12. Russell, K.P. : *In Training Examination for Resident in Obstetrics and Gynecology, Report of the 1970 Examination*. Obstet. Gynecol., 36-6: 953-956, 1970.