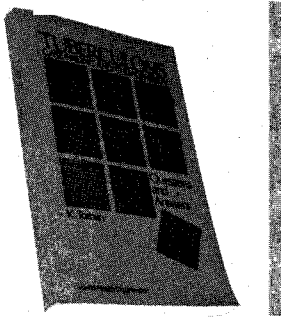


결핵의 발견과 화학요법

항결핵화학요법 29

권 동 원 / 본회 역학부장, 결핵전문이



이글은 WHO가 발행한 토만저 「결핵의 발견과 화학요법」을 번역한 글이다.

측정자 오차(observer error)

흉부엑스선검사의 효과는 비정상적인 음영과 그들을 정확히 해석할 수 있는 판독자의 능력에 의하여 주로 결정된다. 이 말은 판독자가 그 비정상적인 음영을 발견하지 못하거나 과소판독(under-reading)하지 않는 것을 의미하며 또한 역으로 사진상에 있는 정상적인 음영을 과잉판독(over-reading)하지 않는 것을 의미하는 것이다. 이 능력은 판독자

마다 차이가 있다(개인간의 차이: inter-individual variation). 그러나 동일한 판독자라도 초회판독시에는 관찰하였던 어떤 비정상적 음영을 일주일 정도 시간이 경과한 후에 동일한 사진을 다시 판독하면 관찰하지 못하는 경우도 또한 있다. 이와 반대로 초회판독에서는 정상으로 여겼던 비정상적 음영을 재판독시에 발견하는 수도 있다. 이 모순(개인의 차이: intraindividual variation)은 판독자간의 차이보다 더 격정된다.

측정자 오차는 흉부엑스선검사에만 있는 현상이 아니다. 매일 사용되며 또한 정확하고 객관적인 것으로 생각되는 많은 임상검사와 실험절차들, 예를들면, 혈압측정, 심전도검사, 혈구수측정, 내시경검사, 비색정량법(visual colorimetric methods), 양피부검사(quantitative skin tests)와 같은 것에도 측정자 오차가 있다. 측정자 오차의 정도를 알기 위한 연구가 수십년전 특히 세계 제2차대전후에 많은 선진공업국가에서 항결핵운동이 시작되었을 때 있었다. 실제적으로 첫 연구들은 다양한 엑스선 및 형광엑스선 투시(photofluorographic)기법과 작은 거

을 카메라 필름과 같은 기구들의 효과를 알기 위해 실시 되었다. 초기 연구들이 거의 대부분은 Yerushalmy에 의해서 고안되거나 실시되었는데 그는 생통계학자로서 그가 발견한 것들은 방사선의사와 흉곽의사의 관심을 곧 불러 일으켰다.

과잉판독과 과소판독. 한 실험에서는, 5명의 전문가들이 필름의 다양한 크기가 흉부엑스선검사 결과에 미치는 영향에 관하여 조사하였다. 그들은 필름의 크기가 측정자 오차(observer variation)의 정도보다 훨씬 덜 중요하다는 것을 발견하였다. 각각의 전문가들은 매 시리즈마다 '양성' 사진의 약 25%를 발견하지

(표1)

측정자 오차 : 흉부엑스선 사진의 과소판독 및 과잉판독
(대부분 비선택 조사 자료)

연구	과소판독	과잉판독
	양성%	음성%
1. 5명의 전문판독자	25	—
2. 다양한 경험을 소유한 판독자	27	1.7
3. 집단엑스선 검진	32	1.7
4. 덴마크의 결핵중앙등록소, 집단엑스선검진	32	1.6
5. 판독위원 (15,000명의 대학생에 대한 집단엑스선 검진 사진당 10회 판독		
(a) 모든 50명의 판독자	39	1.2
(b) 선택된 판독자		
그룹 A	21	0.5
그룹 B	26	0.3

못했다. (과소판독 하였다)(표1). 동일한 사진을 약 3개월후에 다시 판독하였을때 각 판독자는 약 1/5에서 마음을 바꾸어 첫 판독시에 '양성'으로 분류했던 것을 반복하였다(약 20%의 개인의 차이). 이러한 발견에 놀라서, 두 그룹의 방사선의사들이 이 문제를 더 조사 하였다. 그들의 결론은 그 발견이 정말로 맞다는 것이었다.

후속 연구들에 의해서 26-43%의 사진이 과소판독 된다는 것이 확인 되었다. 따라서 3명의 경험있는 형광엑스선 사진판독자들로 구성된 한 덴마크 그룹이 5,000매의 비선택 간찰사진을 독립적으로 판독하였는데(표1), 평균 32%의 과소판독과 약 2%의 과잉판독이 있는 것이 발견 되었다. 이러한 사실은 영국에서 Cochrane에 의하여 확인되었다.

정기적인 엑스선검사의 유용성을 알기 위한 한 미국의 연구에서 15,000명의 대학교 1학년 학생이 형광엑스선 사진을 찍었다. 그리고는 동수의 방사선의사와 흉곽전문의들로 구성된 50명의 판독위원들이 사진을 판독 하였다. 무작위로, 각 판독자는 3,000장의 사진을 판독하여 각 사진을 독립적으로 10번 판독하도록 하였다. 한사람 혹은 그 이상의 판독자에 의해서 사진이 양성으로 판독된 모든 학생들은 세균학적검사, 투베르쿨린 피부반응검사 및 단층촬영을 시행하고 그 학생들이 대학교에 재학하는 동안 추구조사를 하였다. 결국은 249장의 사진이 모든 필요한 정보를 소유하고 있는 소 그룹의 최종판독자(umpire reader)들에 의하여 '확실한 양성'으로 분류 되었다. 개 인간의 차이의 정도가 상당하였다. 50명

의 판독위원 전체의 과소판독의 정도는 양성으로 동의된 249장의 사진 중 평균 39%이었고, 1.2%(156장의 사진)는 과잉판독 되었다.

10명의 '최고' 판독자들(5명의 방사선 의사와 5명의 흉곽전문)의 결과만을 보면 과소판독률과 과잉판독률이 다소 낮기는 하지만 여전히 만족스럽지 못하게 높다(표1, 연구 5a와 5b).

경험이 엑스선사진 판독결과에 미치는 영향. 도표에 있는 결핵연구소에서 좀더 최근에 실시한 연구에서는 일본의 전국적인 환자발견사업에 참여한 192명의 의사들에 의한 과잉판독과 과소판독의 정도를 조사하였다. 엑스선사진 판독 경험이 판독자 차이의 정도에 미치는 영향에 대하여 특별한 주의를 기우렸다(표 2).

건강상태가 결핵연구소에 잘 알려져 있는 50명의 엑스선사진이 독립적 판독을 위하여 선택되었다. (표2)

25장의 사진은 결핵이나 혹은 다른 폐질환으로 확진된 사람들의 사진이고, 5장은 결핵이 완치된 사람들의 사진이며, 20장은 어떤 비정상적인 음영도 가지고 있지 않는 사람의 사진이었다. 판독자들에게는 후속검사가 필요한지 안한지 만을 결정하도록 요구했다. 판독자가 비정상적인 음영을 가지고 있는 사람에게 후속검사를 요청하지 않았으면 과소판독한 것으로

**국제항결핵연맹은
흉부엑스선사진 판독과 해석에
관한 가장 중요한 비교연구의
하나를 준비하였다. 이 연구의
주된 목적은 흉부엑스선사진의
국제적 분류의 기본으로 사용할 수
있는 엑스선 음영의 통일된 명명과
해석을 도출해 내는 것이었다.**

기록하였다. 정상 흉부엑스선사진 소견을 가지고 있는 사람에게 후속검사를 요청한 경우에는 과잉판독으로 간주하였다. 비교된 판독자들은 1년 미만에서 10년 이상의 실무경험을 가지고 있었고 연간 1,000장에서 20,000장 혹은 그 이상의 사진을 판독하였다.

평균 과소판독률은 21.8%이었고 과잉판독률은 19.5%이었다. 그러나 10년 이상의 경험을 가지고 있거나 연간 20,000장 이상의 사진을 판독하였던 판독자들은 다른 판독자들 보다 과소판독률이 약

측정자 오차 : 흉부엑스선사진

(70mm거울반사카메라)의 과소판독 및 과잉판독. 일본의 국가결핵관리사업중 환자발견사업에 참여한 192명의 의사들의 판독결과에 경험이 미치는 영향. (일본 항결핵협회, 결핵연구소의 연구)

경험	판독자수	과소판독	과잉판독
		%	%
(a) 1-4년	37	28.0	18.0
5-9년	37	19.2	19.0
>10년	88	17.6	17.0
혹은			
(b) 연간 1-5,000장의 사진	43	22.4	17.5
연간 5,000-20,000장의 사진	48	24.0	18.0
>연간 20,000장의 사진	41	15.2	15.5
전체 판독자의 평균	192	21.8	19.5

6-8%가 낮았다. 최소한 2번의 판독범실을 저지르지 않았던 판독자는 한사람도 없었다. 연구자들은 일본에서 실시되는 집단엑스선검사에서 활동성 결핵환자의 약 1/5을 아마도 발견하지 못하고 있을 것이라고 추정하였다. 집단엑스선검사에 쏟는 상당한 노력과 이 분야에 있어서 일본의 높은 기술수준을 고려해 볼 때 그것은 놀랄만한 발견이었다.

경과관찰(follow-up)을 위한 흉부엑스선사진 판독에서의 의견 불일치, 판독자간의 의견 불일치는 환자발견 및 진단 목적으로 엑스선사진을 판독할 때만 있는 것이 아니라 이미 진단된 환자의 연속사진들을 경과관찰을 위하여 비교판독할 때에도 있다. 따라서 한 연구에서 한 환자당 서로 다른 시기에 촬영한 2장의 사진(모두 9,000쌍의 사진)을 판독하였다. 판독자들에게는 두번째의 사진이 첫번째 사진보다 호전되었는가 악화되었는가 혹은 변화가 없는가를 보고하도록 요구 했다. 결과는 표3에 나타나 있다.

사진들이 3명의 방사선의사와 3명의

(표3)
판독자 차이 : 흉부엑스선사진 해석에 있어서의 의견불일치
(35.6×43cm)

판독자	개인간의	개인의
	의견불일치	의견불일치
	%	%
(a) 2 그룹의 전문가 (3명의 방사선의사와 3명의 흉곽전문의)		
그룹 A	29	19
그룹 B	27	24
(b) 2명의 전문판독자(동일한 자료를 판독함)	30	21

흉곽전문의들로 구성된 두 그룹에 의하여 판독되었건 혹은 2명의 전문 판독자에 의해서만 판독되었건 간에 결과는 크게 다르지 않았다. 판독자간의 의견 불일치 수준은 약 27-30%이었고 판독자의 의견불일치 수준은 약 19-24%이었다.

국제항결핵연맹(IUAT)의 엑스선 분류에 관한 국제적 연구

국제항결핵연맹은 흉부엑스선사진 판독과 해석에 관한 가장 중요한 비교연구의 하나를 준비하였다. 이 연구의 주된 목적은 흉부엑스선사진의 국제적 분류의 기본으로 사용할 수 있는 엑스선 음영의 통일된 명명과 해석을 도출해 내는 것이었다. 노르웨이의 상인인구 집단을 주기적으로 집단검진 하면서 촬영한 수십만장의 사진중 1,100장을 표본으로 택하였다. 표본에는 전염성 결핵환자의 사진이 200장, 한 때 활동성 폐결핵환자였던 사람의 사진

이 400장, 후송이나 추구조사를 필요로 하지 않는 경증환자의 사진이 100장, 비정상적 음영이 없는 사람의 사진이 300장, 비결핵성 폐질환으로 확인된 사람의 사진이 100장이 포함되어 있었다. 그 사진들을 7개의 사진 롤에 감아서 각 롤을 10개씩 복사하였다.

이 일에 굉장히 유능한 90명의 의사(방사선의사와 흉곽의사가)가 판독을 하

였다. 80명의 판독자들은 수년동안 집단 엑스선검진사업을 실시해 오고있는 9개의 나라 즉 체코슬로바키아, 덴마크, 필란드, 폴란드, 노르웨이, 스웨덴, 영국, 미국, 유고슬라비아에서 선택되었고, 10명의 판독자들은 다양한 국가적 및 국제적 WHO사업 참모 중에서 선택되었다.

이 연구는 주로 판독자 사이의 의견 일치 혹은 불일치의 정도를 주로 측정하고자 고안된 것이지, 과소판독 혹은 과잉판독을 일으키는 판찰자 오차를 측정하고자 고안된 것은 아니다. 미리 준비된 한 세트의 질문에 대하여 각 판독자들은 독립적으로 대답하였다. 대부분의 질문들은 “폐에 어떤 비정상적인 음영이 있습니까? 예/아니오?”, “공동이 있습니까? 예/아니오?”, “환자는 임상적인 주의를 필요로 합니까? 예/아니오?” 등과 같이 예 혹은 아니오라는 답을 필요로 하였다.

1, 100장의 사진 하나하나에 대한 모든 질문이 90명의 판독자들에 의해서 대답되어야 하므로 각 질문에 대한 답은 99,000개가 있었다. 엄청난 양의 수집된 자료는 전자 자료처리기(electronic data processing machinery)에 의해서 분석되고 도표가 작성되었다.

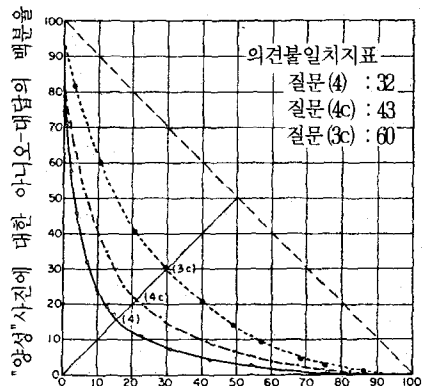
이 자료는 일련의 값이 각 질문에 대하여 얻어지는 특별한 통계학적 절차에 의하여 평가되었다. 이 값들로부터 어떤 주어진 질문에 대한 판독자간의 의견 일치(혹은 오히려 불일치)수준을 나타내는 곡선을 그렸다(그림1) 이 방법의 장점을 이러한 방법으로 얻어진 곡선에 한 수치 즉 한 지표로 나타낼 수 있다는 것이다. 이 지표는 0에서 100까지의 수를 취하여

의견불일치의 정도를 반영할 수 있다. 따라서 낮은 지표는 의견불일치가 적은 것을 의미하고 높은 지표는 의견불일치가 많다는 것을 의미하는 것이다. 곡선이 두축의 0점에 가까울 수록 의견불일치율은 낮고 곡선이 평편할 수록 혹은 축에서 멀어질 수록 의견불일치율이 높은 것이다. (100이라는 최대치는 왼쪽 윗 모퉁이에서 오른쪽 아래 모퉁이로 그 중간에 꿩건선과 일치하고 있다.)

예를들면, 질문4에 대한 것이 질문 3C에 대한 것보다 의견불일치율이 낮은 것이다(그림1). (이 지표는 곡선상에 있는 백분율이 동일한 지점, 즉 곡선이 0 점 모퉁이로부터 시작해서 사선으로 그은 선과 교차되는 곳에서 서로 다른 답의 100분율을 합함으로 계산된다.)

(그림 1)

IUAT의 엑스선 분류에 관한 연구: 3개의 질문(4: 호흡기에 어떤 이상이 있는가?, 4c: 림프절에 어떤 이상이 있는가? 3c: 림프절에 어떤 석회화 된 것이 있는가?)에 대한 판독자간의 의견불일치 곡선



"음성" 사진에 대한 예-대답의 백분율