

유방촬영술의 임상영상검사: 신 임상영상평가표와 신 임상영상평가표의 개정안의 비교

신희정¹ · 김학희¹ · 손은주² · 서보경³ · 고은영⁴ · 고경희⁵
김대봉⁶ · 백문희⁷ · 류정아⁸ · 최연현⁴

¹울산의대 서울아산병원 영상의학과, 영상의학연구소

²연세의대 영동세브란스병원 영상의학과

³고려대학교의료원 안산병원 영상의학과

⁴성균관대의대 삼성서울병원 영상의학과

⁵포천중문의대 분당차병원 영상의학과

⁶분당재생병원 영상의학과

⁷포천중문의대 강남차병원 영상의학과

⁸관동의대 명지병원 영상의학과

목적: 2002년 제정된 신 임상영상평가표와 2006년 신 임상영상평가표에 감점항목을 추가하여 보완한 신 임상영상평가표의 개정안의 정확도와 객관성을 비교하고자 하였다.

대상 및 방법: 다양한 의료기관에서 시행한 유방촬영술 72건을 유방영상의학을 전공으로 하며, 4년 이상의 경험을 가진 4명의 영상의학과 전문의와 3년 이하의 경험을 가진 3명의 영상의학과 전문의가 각각 평가하였다. 평가방법은 우선 각각의 유방촬영술에 대하여 주관적 기준으로 보았을 때 합격, 불합격 여부를 평가한 후, 평가표에 따라 채점을 하였다. 신 임상영상평가표로 먼저 채점하고 1개월 이후 신 임상영상평가표의 개정안으로 채점하도록 하였다. 총점 100점 만점에 60점 이상을 합격으로 판정하였다. 또한 4년 이상의 경험을 가진 4명의 영상의학과 전문의에서 각 항목 별로 점수 변화를 알아보았다.

결과: 주관적 기준으로 볼 때 72건 중 31건(43%)이 불합격으로 판정되었고, 평가표에 의한 불합격률은 신 임상영상평가표에서 9 - 26건(13 - 36%)으로 평균 22%이었고, 신 임상영상평가표의 개정안에서는 9 - 40건(13 - 56%)으로 평균 30%이었다. 주관적 평가와 평가표와의 일치도는 신 임상영상평가표의 개정안이 신 임상영상평가표에 비하여 더 큰 일치도를 보였다. 주관적 평가에 의해 불합격한 유방촬영술을 점수로 계산한 평가표로 불합격으로 판독하는 민감도는 신 임상영상평가표에서 평균 42%였고, 신 임상영상평가표의 개정안에서 평균 60%로 증가하였다. 주관적 평가에 의해 합격한 유방촬영술을 점수로 계산한 평가표로 합격으로 판독하는 특이도는 신 임상영상평가표에서 평균 92%였고, 신 임상영상평가표의 개정안에서 평균 93%로 비슷하였다.

결론: 신 임상영상평가표의 개정안이 신 임상영상평가표에 비하여 유방촬영술의 임상영상평가에 더 정확하고 유용하다.

Index words: Mammography, Quality assurance

통신저자 : 김학희

(138 - 376) 서울시 송파구 풍납2동 388 - 1, 울산의대 서울아산병원 영상의학과, 영상의학연구소

Tel. (02)3010 - 4390, Fax. (02)476 - 0090

E - mail: hhkim@amc.seoul.kr

서 론

우리나라에서 유방암 발생률은 매년 증가하고 있으며, 2002년 이후 위암을 앞질러 여성암 발생 1위를 차지하고 있다 (1). 미국과 유럽 등 선진국에서는 지난 20여년간 유방암 조기 발견을 위한 선별 유방촬영술을 국가적 차원에서 시행하고 있으며, 선별 유방촬영술을 통해 유방암으로 인한 여성 사망률을 감소시키고자 노력하고 있다. 최근 우리나라에서도 특정암 검진사업에 유방촬영술을 이용한 유방암 검진이 포함되었고, 이로 인해 여성들의 선별 유방촬영술 검사건수도 급격히 증가하고 있다. 그러나 우리나라 유방촬영기의 품질관리가 미흡하여 화질관리가 제대로 되지 않아 조기 유방암을 놓치거나 오진할 가능성이 많았으며, 실제로 이러한 사례가 법정 소송으로 이어지는 등 사회 문제화 되고 있다 (2, 3). 우리나라 여성들의 유방은 구미 여성들에 비해 치밀한 밀도로 보이는 경우가 많기 때문에 유방암의 발견이 용이하지 않을 수 있다 (4). 따라서 양질의 유방촬영화질과 정확한 진단을 위해 유방촬영술에 대한 체계적인 관리가 필요하다.

미국과 유럽 등 선진국에서는 일찍부터 유방촬영기의 화질관리에 대한 자발적 또는 강제적 인증제도가 정착화 되어 있다 (5). 미국에서는 1980년대 초부터 방사선의 학회와 식품의약품관리국의 협조아래 유방촬영술 시행기관에 대한 화질관리 인증제도 프로그램을 시행하였고, 1992년 유방촬영술 정도 관리 표준법(Mammography Quality Standards Act)을 제정하여 유방촬영술의 화질관리를 법적으로 관리하고 있다 (6, 7). 우리나라에서도 1995년 진단용 방사선발생장치의 안전관리에 관한 규칙이 제정되어 영상진단검사의 정도 관리에 관한 지침을 마련하였고, 대한영상의학회와 대한유방영상의학회에서 유방촬영술의 정도 관리에 대한 필요성을 제기해 왔다. 2003년 1월 특수의료장비의 설치 및 운영에 관한 규칙이 제정 공포됨에 따라 유방촬영용장치의 품질관리 중 임상영상검사의 평가는 3년마다 시행하는 정밀검사 시 시행하게 되었다 (8). 그러나 최초의 임상영상평가표는 기준이 다소 애매하여 관찰자 내 또는 관찰자 간 일치도가 떨어질 우려가 있고, 정확한 평가가 어려울 수 있으며, 이에 대한 몇몇 연구들이 있었다 (9 - 11). 이에 2002년 대한유방영상의학회 정도관리위원회에서 신임상영상평가표를 제시하였고, 이 신평가표로 좀더 객관적인 평가가 가능해졌으나, 평가의 특이도가 다소 낮은 단점이 있었다. 이런 단점을 보완하고자 2006년 신평가표에 감점항목을 추가하였고, 항목 중 일부를 세분화하여 개정

안을 만들게 되었다. 본 연구는 2002년 대한유방영상의학회 정도관리위원회에서 제시한 신 임상영상평가표(이하 신평가표)와 2006년 신평가표에 감점항목을 추가하여 새롭게 만든 신평가표의 개정안(이하 신평가표의 개정안)의 정확도와 객관성을 비교하고자 하였다.

대상 및 방법

2005년 본원에 내원한 환자들 중 다른 병원에서 촬영한 유방촬영술이 있었던 72건을 대상으로 하였다. 다른 병원에서 촬영한 유방촬영술은 2003년 3월부터 2005년 2월에 촬영된 것이었으며, 복사필름은 포함시키지 않았다. 이 중 66건은 2005년 백문희 등(11)이 시행한 연구에서와 같은 유방촬영술을 대상으로 하였고, 나머지 6건은 필름 분실로 인하여 새로운 유방촬영술을 대상으로 하였다. 46개의 병, 의원에서 촬영된 것으로 3건은 대학병원(4%), 69건은 일차의료기관(96%)에서 시행한 것이었다. 2002년 대한유방영상의학회 정도관리위원회에서 제시한 신평가표 (Table 1)와 2006년 신평가표에 감점항목을 추가하여 새롭게 만든 신평가표의 개정안 (Table 2)를 이용하여 유방영상의학을 전공으로 하는 4년 이상의 경험을 가진 4명의 유방영상의학과 전문의와 3년 이하의 경험을 가진 3명의 유방 영상의학과 전문의가 각각 평가하였다. 평가방법은 7명의 전문의가 주관적 기준으로 보았을 때 합격, 불합격 여부를 평가한 후, 두 개의 평가표를 가지고 각각 채점을 하였다. 이 때 신평가표로 먼저 채점하고, 1개월 이후 신평가표의 개정안으로 채점하도록 하였다. 각 검사자는 채점표의 총점과 합격 여부를 알지 못하도록 하였고, 합격기준은 총점 100점 만점에 60점 이상인 경우로 정하였다. 주관적 평가를 기준으로 한 각 평가표의 민감도와 특이도를 paired t-test를 이용하여 분석하였고, 주관적 평가와 비교한 신평가표 또는 신평가표의 개정안의 일치도는 교차분석을 이용하여 kappa value를 구하였다. 신평가표의 개정안에서 감점항목을 추가하고, 점수 배점을 달리하여 강화된 자세잡기, 압박, 대조도와 노출, 및 촬영표지 항목에 대하여는 4년 이상의 경험을 가진 4명의 유방영상의학과 전문의에서 채점한 점수를 3개의 항목 별로 신평가표와 신평가표의 개정안간 차이가 있는지 비교하였고, 항목간 비교는 ANOVA test를 사용하여 통계적 분석을 하였다.

Table 1. 신 임상영상평가표

일련번호	촬영병원	환자이름	환자등록번호	촬영일	평가일	평가자
			200 . . .	200 . . .		
항목	평가 내용					점수표시
1.촬영 표지	1. 환자의 이름 2. 성별 3. 나이 4. 환자(병록) 번호 5. 촬영날짜 6. 촬영기관 이름 7. 내외사위 또는 상하 영상, 좌우와 외측 표지 8. 카세트 번호 9. 촬영자 이름(기호) 10. 필름에 직접 표지하지 않고 종이 등에 써서 부착하였을 때					1 1 1 1 1 1 1 1 1 - 4 (감점)
2.자세 잡기 내외사위 영상	1. 대흉근이 잘 보이는가? • 대칭 역삼각형으로 외연이 볼록하거나 직선으로 보인다 • 대흉근이 오목하게 보인다. • 대흉근이 일부만 포함되었다. 2. 대흉근의 하단이 유두 후방선 보다 같거나 더 아래로 보이는가? • 같거나 더 아래로 나온 상태로 보인다 • 대흉근의 하단이 유두 후방선보다 위에 있다. 3. 유방의 아래 부위의 처짐(sagging)이 있는가? (대흉근이 일부만 또는 전혀 포함 되지 않은 경우는 0점 처리) • 팽팽하다. • 약간 처져 있다 • 많이 처져 있다. 4. 유방하 주름(inframammary fold)이 포함되었는가? • 보인다 (포함되었다). • 보이지 않는다. 5. 신체의 다른 부분이 유방을 가리는가? • 유방의 일부를 전혀 가리지 않는다 • 유방의 일부를 약간 가린다. • 유방이 많이 가려 졌다. 6. 피부의 겹침과 주름은? • 전혀 없다. • 약간 있다 • 판독에 지장을 줄 정도로 많이 보인다. 7. 유두 • 측면상으로 유두의 윤곽 전체가 관찰된다. • 유방 조직과 약간 겹쳐 있다. • 유방 조직 속에 보인다.					5 3 0 5 0 5 3 0 5 0 3 1 0 3 1 0 3 1 0
상하 영상	8. 유두 • 유방의 중앙에 측면상으로 유두 윤곽 전체가 관찰된다. • 유방 조직과 약간 겹쳐 있다. • 유방 조직 속에 보인다. 9. 유두후방선의 길이를 내외사위 영상과 비교하면? • 차이가 1 cm 이내이다. • 차이가 1-2 cm 이다. • 차이가 2 cm 이상이다. 10. 내측 유방조직이 포함된 정도는? • 충분히 포함되어 있다. • 포함되지 않은 부분이 상당히 많다.					3 1 0 5 3 0 4 0

	11. 유방 후방 지방(retromammary fat)이 포함되었는가? • 많이 포함 되었다. • 약간 포함되었다. • 전혀 포함되지 않았다. 12. 피부의 접침과 주름은? • 전혀 없다. • 약간 있다 • 판독에 지장을 줄 정도로 많이 보인다.	5 3 0 3 1 0
3. 압박	1. 압박 정도는? • 충분히 압박되어 유방조직이 분리되어 보인다. • 불충분하게 압박되어 유방조직이 많이 중첩되어 있다. 2. 움직임이 있는가(motion artifact)? • 움직임에 의한 영상의 흐림이 없다.. • 움직임에 의한 영상의 흐림이 있다.	4 0 6 0
4. 대조도와 노출	1. 지방 조직과 유선 조직이 구분되는가(대조도) ? • 뚜렷이 구분되어 보인다. • 구별은 가능하나 선명하지 않다. • 구별이 어렵다. 2. 노출이 판독하기에 적절한가? • 별도의 밝은 전등을 사용하지 않아도 판독이 가능하다(과다 노출이 없다). • 별도의 밝은 전등을 사용하면 유방 조직이 잘 보인다. • 유선이 지나치게 밝게 보이며 피부선이 쉽게 잘 보인다(노출 부족).	4 2 0 6 3 0
5. 잡영과 인공물	1. 점상 인공물 • 5 개 이상 • 1-4개 • 없다 2. 굵힘 • 2 개 이상 • 1 개 • 없다 3. 지문 • 2 개 이상 • 1 개 • 없다 4. 롤러 자국 • 2 개 이상 • 1 개 • 없다 5. 그리드 선이 보이는가? • 보인다. • 안 보인다. 6. 현상 과정 전후에 발생한 안개(fog) 모양의 인공물 • 2 개 이상 • 1 개 • 없다 7. 증감지-필름 접촉 불량에 의한 인공물 • 2 개 이상 • 1 개 • 없다. 8. 기타 인공물 • 2 개 이상 • 1 개 • 없다.	0 1 2 0 1 2 0 1 2 0 1 2 0 2 0 1 2 0 1 2 0 1 2

6. 기타 사항	1. 필름의 유방 조직 바깥 부분이 검게 유지되었는가? -시준기(collimator) 사용 여부	
	• 필름의 바깥 부분이 검게 남아 있다.	3
	• 필름의 바깥 부분이 희게 촬영되어 있다.	0
	2. 필름을 적절하게 사용하였는가?	
	• 적절한 크기의 필름을 사용하였으며, 한 필름에 한 영상을 촬영하였다.	3
	• 유방의 크기에 비해 필름이 작거나 한 필름에 두 영상을 촬영하였다.	0
총점		

결 과

주관적 기준을 기준으로 하여 72건 중 41건(57%)을 합격으로, 31건(43%)을 불합격으로 판정하였다. 평가표에 의한 불합격률은 신 평가표에서 9-26건으로 불합격률은 평균 22%(범위, 13-36%)이었고, 신 평가표의 개정안에서는 9-40건으로 평균 30%(범위, 13-56%)이었다.

교차분석에 의하여 주관적 평가를 기준으로 한 신 평가표와 신 평가표의 개정안의 합격과 불합격의 일치도를 구하였다. 7명 검사자의 kappa value는 신 평가표에서 0.228에서 0.567로 평균 0.357이었고, 신 평가표의 개정안에서 0.318에서 0.770으로 평균 0.546이었다. 7명 검사자 중 5명에서 kappa value가 증가하였고, 1명에서는 변화없었으며, 나머지 1명에서는 0.03정도 감소하였다. 그러나 신 평가표와 신 평가표의 개정안간에 통계적 유의한 차이를 보였고, 신 평가표의 개정안이 신 평가표보다 주관적 평가에 대한 일치도가 더 높았다.

Paired t-test를 이용하여 주관적 평가를 기준으로 하여 불합격한 유방촬영술을 평가표로 불합격으로 판독하는 민감도는 신 평가표에서 26-68%로 평균 42%였고, 신 평가표의 개정안에서 29-81%로 평균 60%였다. 또한, 주관적 평가를 기준으로 하여 합격한 유방촬영술을

평가표로 합격으로 판독하는 특이도는 신 평가표에서 85-98%로 평균 92%였고, 신 평가표의 개정안에서 68-100%로 평균 93%였다.

ANOVA test로 신 평가표의 개정안에서 감점 항목을 추가한 자세잡기와 촬영표지 항목과 점수를 강화한 압박, 대조도와 노출 항목에 대하여 항목간 비교를 하였다. 4년 이상의 경험을 가진 4명의 영상의학과 전문의에서 항목간 비교를 한 결과 자세잡기 항목에서 2명의 검사자에서 점수가 감소하였고, 나머지 2명의 검사자에서는 점수가 증가하였으며, 두 개의 평가표간 통계적으로 유의한 차이

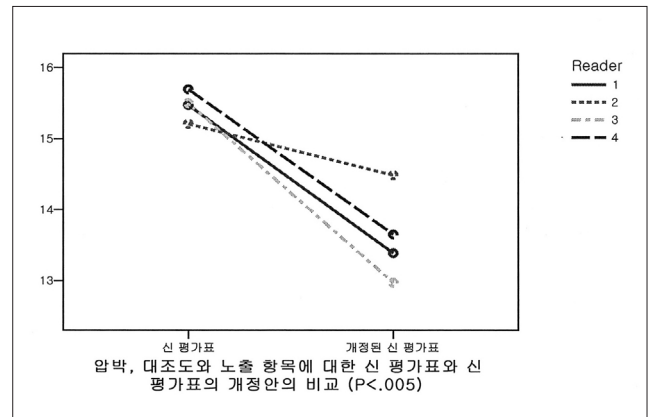


Fig. 2. 압박, 대조도와 노출 항목에 대한 신 평가표와 신 평가표의 개정안의 비교

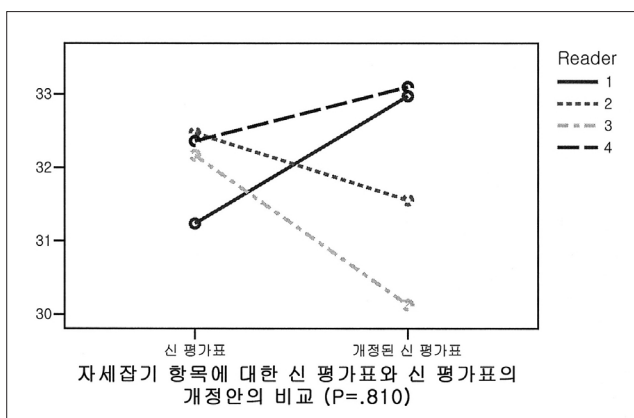


Fig. 1. 자세잡기 항목에 대한 신 평가표와 신 평가표의 개정안의 비교 (Y축; 자세잡기 항목의 평균 점수)

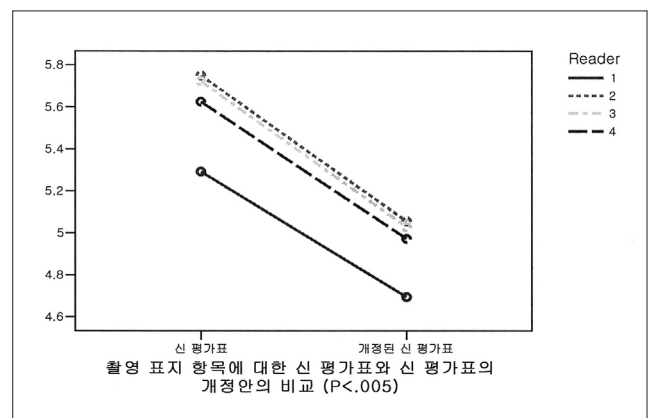


Fig. 3. 촬영 표지 항목에 대한 신 평가표와 신 평가표의 개정안의 비교

Table 2. 신 임상영상평가표의 개정안

일련번호	촬영병원	환자이름	환자등록번호	촬영일	평가일	평가자
			200 . . .	200 . . .		
항목	평가 내용					점수표시
1.촬영 표지	1. 환자의 이름 2. 성별 3. 나이 4. 환자(병록) 번호 5. 촬영날짜 6. 촬영기관 이름 7. 내외사위 또는 상하 영상, 좌우와 외측 표지 8. 카세트 번호 9. 촬영자 이름(기호) 10. 필름에 직접 표지하지 않고 종이 등에 써서 부착하였을 때 환자 인적사항이 표기되었으나 알아볼 수 없는 경우					1 1 1 1 1 1 1 1 1 - 4 (감점)
2.자세 잡기 내외사위 영상	1. 대흉근이 잘 보이는가? • 대칭 역삼각형으로 외연이 불룩하거나 직선으로 보인다 • 대흉근이 오목하게 보인다. • 대흉근이 일부만 포함되었다. • 대흉근이 전혀 포함 되지 않았다. 2. 대흉근의 하단이 유두 후방선 보다 같거나 더 아래로 보이는가? • 같거나 더 아래로 나온 상태로 보인다 • 대흉근의 하단이 유두 후방선보다 위에 있다. • 대흉근 하단이 보이지 않거나 너무 위에 위치해 있다. 3. 유방의 아래 부위의 처짐(sagging)이 있는가? (대흉근이 일부만 또는 전혀 포함 되지 않은 경우는 0점 처리) • 팽팽하다. • 약간 처져 있다 • 많이 처져 있다. 4. 유방하 주름(inframammary fold)이 포함되었는가? • 포함되어 아래 조직이 잘 보인다. • 일부 포함 되어 보인다. • 보이지 않는다. 5. 신체의 다른 부분이 유방을 가리는가? • 유방의 일부를 전혀 가리지 않는다 • 유방의 일부를 약간 가린다. • 유방이 많이 가려 졌다. 6. 피부의 겹침과 주름은? • 전혀 없다. • 약간 있다 • 판독에 지장을 줄 정도로 많이 보인다. 7. 유두 • 측면상으로 유두의 윤곽 전체가 관찰된다. • 유방 조직과 약간 겹쳐 있다. • 유방 조직 속에 보인다. 8. 유방의 일부 조직이 영상 내에 포함 되지 않았다.					5 3 0 - 5 (감점) 5 3 0 5 3 0 5 3 0 3 1 0 3 1 0 3 1 0 - 5 (감점)
상하 영상	9. 유두 • 유방의 중앙에 측면상으로 유두 윤곽 전체가 관찰된다. • 유방 조직과 약간 겹쳐 있다. • 유방 조직 속에 보인다. 10. 유두후방선의 길이를 내외사위 영상과 비교하면? • 차이가 1 cm 이내이다. • 차이가 1-2 cm 이다. • 차이가 2 cm 이상이다.					3 1 0 5 3 0

	11. 내측 유방조직이 포함된 정도는? • 충분히 포함되어 있다. 4 • 포함되지 않은 부분이 상당히 많다. 0 12. 유방 후방 지방(retromammary fat)이 포함되었는가? • 많이 포함 되었다. 5 • 약간 포함되었다. 3 • 전혀 포함되지 않았다. 0 13. 피부의 접침과 주름은? • 전혀 없다. 3 • 약간 있다 1 • 판독에 지장을 줄 정도로 많이 보인다. 0 14. 유방의 일부 조직이 영상 내에 포함 되지 않았다. - 5 (감점)	
3. 압박	1. 압박 정도는? • 충분히 압박되어 유방조직이 분리되어 보인다. 4 • 압박이 덜 되어 일부 유방조직이 겹쳐 있다. 2 • 불충분하게 압박되어 유방조직이 많이 중첩되어 있다. 0 2. 움직임이 있는가(motion artifact)? • 움직임에 의한 영상의 흐림이 없다. 6 • 움직임에 의한 영상의 흐림이 약간 있다. 3 • 움직임에 의한 영상의 흐림이 많다. 0	
4. 대조도와 노출	1. 지방 조직과 유선 조직이 구분되는가(대조도)? • 뚜렷이 구분되어 보인다. 6 • 구별은 가능하나 선명하지 않다. 3 • 구별이 어렵다. 0 2. 노출이 판독하기에 적절한가? • 별도의 밝은 전등을 사용하지 않아도 판독이 가능하다(과다 노출이 없다). 6 • 별도의 밝은 전등을 사용하면 유방 조직이 잘 보인다. 3 • 유선이 지나치게 밝게 보이며 피부선이 쉽게 잘 보인다(노출 부족). 0	
5. 잡영과 인공물	1. 점상 • 5 개 이상 0 • 1-4개 1 • 없다 2 2. 굵힘, 지문 • 2 개 이상 0 • 1 개 1 • 없다 2 3. 롤러 자국 • 2 개 이상 0 • 1 개 1 • 없다 2 4. 그리드 선이 보이는가? • 보인다. 0 • 안 보인다. 2 5. 현상 과정 전후에 발생한 안개(fog) 모양의 인공물 • 2 개 이상 0 • 1 개 1 • 없다 2 6. 증감지-필름 접촉 불량에 의한 인공물 • 2 개 이상 0 • 1 개 1 • 없다. 2	

	7. 기타 인공물 • 2 개 이상 • 1 개 • 없다. 8. 인공물이 판독에 지장을 주는 경우	0 1 2 - 5(감점)
6. 기타 사항	1. 필름의 유방 조직 바깥 부분이 검게 유지되었는가? -시준기(collimator) 사용 여부 • 필름의 바깥 부분이 검게 남아 있다. • 필름의 바깥 부분이 희게 촬영되어 있다. 2. 필름을 적절하게 사용하였는가? • 적절한 크기의 필름을 사용하였으며, 한 필름에 한 영상을 촬영하였다. • 유방의 크기에 비해 필름이 작거나 한 필름에 두 영상을 촬영하였다.	3 0 3 0
총점		

는 보이지 않았다 ($p = 0.810$) (Fig. 1). 압박, 대조도와 노출, 및 촬영표지 항목에서는 4명의 검사자 모두에서 점수가 감소하였으며, 두 개의 평가표간에 통계적으로 유의한 차이를 보였다 ($p < .005$) (Fig. 2 and 3).

고 찰

2001년 보건복지부의 “방사선검사의 화질 정도 관리를 위한 실태조사” 과제 보고서에서 최초로 전국 유방촬영술 화질에 대한 실태조사와 유방팬텀을 이용한 현장검사 자료를 제시하였다 (12). 이후 2003년 문우경 등은 유방촬영술의 임상영상평가에 대한 전국조사에서 총 598건의 유방촬영 사진을 대상으로 하였으며, 이중 36.3%가 진단적 가치가 없는 것으로 나타났고, 대학병원보다는 비방사선과 의원이나 법인의료단체의 불합격률이 유의하게 높았다는 결론을 내렸다. 또한 이 연구에서 항목별 불합격률은 자세잡기가 24%로 가장 높았고, 표지, 노출, 대조도, 선명도, 압박, 인공물 등의 순이었다 (9). 2005년 이성훈 등은 유방촬영술 정도관리를 위한 현장조사, 팬텀검사 및 임상영상평가에 관한 연구에서 총 371개의 유방촬영술을 대상으로 하였고, 이 중 83.5%에서 적합한 화질을 보였고, 비방사선과 의원 군에서 74.6%로 가장 낮은 적합률을 보였다. 또한, 임상영상평가에 대한 두 검사자의 일치도를 조사하였을 때 상관관계수가 0.93으로 관찰자 간의 판정에 일치도가 높았다 (10). 위 두 연구에서 대학병원과 종합병원이 차지하는 비율은 각각 32%와 57%였으며, 본 연구에서는 4%로 일차의료기관이 차지하는 비율이 96%로 높았다. 본 연구에서 실제로 정도 관리가 필요한 일차의료기관이 평가대상의 대부분을 차지하였다. 평가표에 의한 불합격률은 이전 두 연구에서 각각 36%와 17%였고, 본 연구에서는 신 평가표에서 평균 22%였고, 신 평가표의 개정안에서 평균 30%였다. 두 연구 중 좀더 최근 임상영상평가에 대하여 시행

한 이성훈 등의 연구에서는 대학병원과 종합병원의 비율이 상대적으로 높고, 본 연구에서는 일차의료기관이 차지하는 비율이 높았고, 이러한 차이에 의해 본 연구에서 불합격률이 더 높았던 것으로 보인다.

2006년 백문희 등은 유방촬영술에 대한 기존 임상영상평가표와 신 임상영상평가표의 비교 연구에서 신 임상영상평가표가 기존 임상영상평가표에 비해 더 정확하고 유용하다는 결과를 얻었다 (11). 이 연구 중 저자들은 신 평가표에서 보완점으로 자세잡기 항목 중 내외사위와 상하위 촬영에서 유방의 일부분이 포함되지 않은 경우, 내외사위 촬영에서 대흉근이 포함되지 않은 경우, 인공물이 판독에 지장을 주는 경우 감점항목을 추가하고, 촬영표지 항목에서 환자 인적 사항이 표기되었으나 알아볼 수 없는 경우도 감점으로 포함하는 등 감점항목을 추가하였고, 대조도와 노출 항목의 점수를 강화하는 등 총점의 증가 없이 평가 항목을 추가한 신 평가표의 개정안을 제시하였다 (Table 2). 신 평가표와 신 평가표의 개정안의 비교에 대한 본 연구에서 신 평가표의 개정안이 신 평가표에 비하여 검사자간 일치도가 높았고, 주관적 평가와 비교한 민감도가 42%에서 60%로 증가하였으며, 특이도는 큰 변화가 없었다. 따라서, 신 평가표의 개정안은 유방촬영술의 임상영상평가 시 불합격을 좀더 잘 판별하는 검사법으로 유용하며, 신 평가표에 비하여 검사자간 일치도가 높은 좀더 객관적인 검사법이라고 보인다.

Bassett 등의 임상화질평가 연구에서 자세잡기가 불합격의 가장 큰 결격사항이었고, 그외 노출, 압박, 선명도, 대조도, 촬영표지, 인공물 등의 순이었다 (13). 우리나라에서는 2001년 보건복지부 과제보고서와 2003년 문우경 등의 유방촬영술의 임상영상평가에 대한 전국조사에 따르면 미국에서의 연구와 마찬가지로 자세잡기 항목이 가장 흔한 결격 항목이었고, 그외 촬영표지, 노출, 대조도 등의 순이었다 (9, 12). 따라서, 신 평가표의 개정안

에서는 신 평가표의 자세잡기 항목에 유방조직의 일부가 포함되지 않거나 내외사위영상에서 대흉근이 포함되지 않은 경우 총점(49점)의 변화없이 감점항목을 추가하였다. 본 연구에서 4년 이상의 경험을 가진 4명의 유방영상의학 전문의가 평가한 자세잡기 항목에 대한 두 평가표의 비교에서 평가표간 통계적으로 유의한 차이를 보이지 않았고, 검사자간 점수폭이 커졌다 (Fig. 1). 이는 검사자에게 평가표의 기준에 대한 교육과 임상영상평가에 대한 실습이 좀더 필요하다고 보인다. 또한, 촬영표지 항목에서 환자 인적 사항이 표기되었으나 알아볼 수 없는 경우도 감점 항목을 추가하였고, 압박, 대조도와 노출 항목의 점수는 20점에서 22점으로 증가하였고, 세부사항을 추가하였다. 본 연구에서 촬영표지와 압박, 대조도와 노출 항목에 대한 두 평가표간 비교에서 신 평가표의 개정안에서 4명의 검사자에서 모두 점수가 감소하였고, 평가표간 통계적으로 유의한 차이를 보였다 (Fig. 2 and 3). 그러나 검사자간 점수폭에는 큰 차이가 없었다.

백문희 등(11)의 연구에서 사용한 자료의 일부와 새로운 자료로 시행한 본 연구 결과를 이전 연구에서와 동일한 7명의 검사자에서 이전에 시행한 결과와 비교하여 주관적 평가에 의한 불합격률이 이전에 25%였고, 본 연구에서 43%로 증가하였다. 또한, 신 평가표에 의한 불합격률도 이전 연구에서 27%(13 - 39%)이었고, 본 연구에서 22%(13 - 36%)이었으며, 신평가표에서 7명 검사자의 평균 kappa value도 이전 연구에서 0.590이었고, 본 연구에서 0.357이었다. 이는 두 연구간에 약 18개월 정도 차이가 있고, 자료의 일부가 다르고, 그 사이에 검사자 들의 경험이 증가하였고, 특히 2 - 3회의 임상영상평가에 대한 교육과 실습을 토대로 검사자 들이 임상영상평가에 대해 좀 더 객관적인 기준을 가짐으로써 생긴 차이로 보인다. 또한, 이전 연구에서는 신 평가표에서 총점 100점 만점에 60점 이상인 경우와 모든 항목 중 촬영표지항목을 제외한 나머지 항목의 부분 합이 54점 이상을 모두 만족한 경우로 정한 반면, 본 연구에서는 총점 100점 만점에 60점 이상인 경우로 하였고, 이러한 차

이도 결과에 차이를 준 것으로 보인다.

결론적으로 유방촬영술의 임상영상평가에 있어서 신 평가표의 개정안이 신 평가표에 비하여 더 객관적이고 정확하다는 것을 알 수 있었다. 그러나, 주관적 평가와의 일치도와 검사자간 일치도가 신 평가표에 비하여 증가하였으나 여전히 다소 낮으며, 이에 대하여는 좀 더 객관적인 평가표에 대한 제시와 검사자에 대한 교육과 실습이 필요할 것으로 보인다.

참 고 문 헌

1. 보건복지부 한국중앙암등록 사업본부. 한국중앙암등록사업 연례 보고서 (2000.1-200.12). 서울, 2002
2. 유근영, 노동영, 이은숙. 유방암의조기검진. 대한의사협회지 2002; 45:992-1004
3. 김철중. 유방암촬영기부정확-오진우려. 조선일보, 서울. 2000년 2월 7일
4. Carmen MGD, Halpern EF, Kopans DB, Moy B, Moore RH, Goss PE et al. Mammographic breast density and race. Am J Roentgenol AJR 2007;188:1147-1150
5. Bassett L. The regulation of mammography. Semin Ultrasound CT MR 1996;17:415-423
6. Hendrick RE, Bassett LW, Bostco MA, et al. Mammography Quality Control Manuals. American College of Radiology, Reston, VA, 1999
7. Barnes GT, Hendrick RE. Mammography accreditation and equipment performance. Radiographics 1994;14:129-138
8. 한국의료영상품질관리원. 특수의료장비의 설치 및 운영에 관한 안내. <http://www.kiami.or.kr>
9. 문우경, 김태정, 차주희, 조경수, 최은완, 이유진 등. 유방촬영술의 임상영상평가. 대한방사선의학회지 2003;49:507-511
10. 이성훈, 최연현, 정수영, 김미혜, 김은경, 오기근 등. 유방촬영술 정도관리를 위한 현장조사, 팬텀 검사 및 임상영상평가. 대한영상의학회지 2005;53:117-127
11. 백문희, 김학희, 손은주, 서보경, 고은영, 고경희 등. 유방촬영술의 임상영상검사 평가: 기존 임상영상평가표와 신 임상영상평가표의 비교. 대한유방검진학회지 2006;3:111-120
12. 2001년 보건복지부 "방사선 검사의 화질 정도 관리를 위한 실태조사 (00-P-14-0006)"과제보고서
13. Bassett LW, Farria DM, Bansal S, Farquhar MA, Wilcox PA, Feig SA. Reasons for failure of a mammography unit at clinical image review in the American College of Radiology Mammography Accreditation. Radiology 2000;215:698-702

Clinical Image Evaluation of Mammograms: Comparison of the New Standard Criteria and the Revised New Standard Criteria

Hee Jung Shin, M.D.¹, Hak Hee Kim, M.D.¹, Eun Ju Son, M.D.², Bo Kyung Seo, M.D.³,
Eun Young Ko, M.D.⁴, Kyung Hee Ko, M.D.⁵, Dae Bong Kim, M.D.⁶,
Moon Hee Paik, M.D.⁷, Jeong Ah Ryu, M.D.⁸, Yeon Hyeon Choe, M.D.⁴

¹*Department of Radiology and Research Institute of Radiology,
University of Ulsan College of Medicine, Asan Medical Center*

²*Department of Diagnostic Radiology, Radiological Institute of Radiological Science, Yongsong Severance Hospital,
Yonsei University College of Medicine*

³*Department of Diagnostic Radiology, Ansan Hospital, College of Medicine, Korea University*

⁴*Department of Diagnostic Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine*

⁵*Department of Diagnostic Radiology, Bundang Cha Hospital, Pochon Cha University, College of Medicine*

⁶*Department of Diagnostic Radiology, Bundang Jesaeng General Hospital, College of Medicine*

⁷*Department of Radiology, Kangnam Cha Hospital, Pochon Cha University, College of Medicine*

⁸*Department of Radiology, Myungi Hospital, Kwandong University, College of Medicine*

Purpose: The purpose of this study was to assess the accuracy of the revised new standard criteria as compared with the new standard criteria in evaluating the quality of mammographic images.

Materials and Methods: A total of 72 mammograms were examined following subjective assessment (pass or fail), the revised new standard criteria with strict guideline, and the new standard criteria for clinical image evaluation, by 7 breast radiologists. The guideline of mammograms for pass is over 60 points out of total 100 points.

Results: Thirty one mammograms (43%) of total 72 mammograms failed in the evaluation following subjective assessment. The average 16 (22%) mammograms failed in the evaluation following the new standard criteria. The average 21 (30%) mammograms failed in the evaluation following the revised new standard criteria. Sensitivity about failure in the evaluation of clinical images following the revised new standard criteria is 60% and is higher than that (42%) of the new standard criteria. Specificity about pass in the evaluation of clinical images following the revised new standard criteria is 93% and is similar to that (92%) of the new standard criteria.

Conclusion: The revised new standard criteria is more accurate and useful in evaluating the quality of mammographic images, as compared with the new standard criteria.

Index words: Mammography, Quality assurance

Corresponding author: Hak Hee Kim, M.D.