

Radiology Residents' Performance in Screening Mammography Interpretation<sup>1</sup>영상의학과 전공의들의 선별유방촬영 판독 성취도<sup>1</sup>Eun Hye Lee, MD<sup>1</sup>, Chae Yeon Lyou, MD<sup>2</sup><sup>1</sup>Department of Radiology, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon, Korea<sup>2</sup>Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Korea

**Purpose:** To evaluate radiology residents' performance in screening mammography interpretation and to analyze the factors affecting performance.

**Materials and Methods:** We enrolled 203 residents from 21 institutions and performed mammography interpretation tests. Between the trainee and non-trainee groups, we compared the interpretation score, recall rate, sensitivity, positive predictive value (PPV) and false-positive rate (FPR). We estimated the training effect using the score differences between trainee and non-trainee groups. We analyzed the factors affecting performance between training-effective and non-effective groups.

**Results:** Trainees were superior to non-trainees regarding interpretation score (43.1 vs. 37.1), recall rate (11.0 vs. 15.5%), sensitivity (83.6 vs. 72.0%), PPV (53.0 vs. 32.4%) and FPR (13.5 vs. 25.5). The longer the training period, the better were the interpretation score, recall rate, sensitivity, PPV and FPR ( $\rho = 0.486, -0.375, 0.343, 0.504, -0.446$ , respectively). The training affected an increase by an average of 6 points; however, 31.6% of institutions showed no effect. A difference was noted in the volume of mammography interpretation during a month (594.0 vs. 476.9) and dedication of breast staff (61.5 vs. 0%) between training-effective and non-effective groups.

**Conclusion:** Trainees showed better performance in mammography interpretation compared to non-trainees. Moreover, performance was correlated with the training period. The factors affecting performance were the volume of mammography interpretation and the dedication of the breast staff.

## Index terms

Breast Imaging  
Education  
Mammography  
Screening

Received October 24, 2012; Accepted February 1, 2013

Corresponding author: Eun Hye Lee, MD

Department of Radiology, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, 170 Jomaru-ro, Wonmi-gu, Bucheon 420-767, Korea.

Tel. 82-32-621-5851 Fax. 82-32-621-5018

E-mail: grace@schmc.ac.kr

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.

This work was supported in part by the Soonchunhyang University Research Fund.

## 서론

선별유방촬영과 국가암검진은 매우 밀접한 관계가 있다. 그 이유는 국가암검진이 본격적으로 시행된 2002년 이후부터 유방암 검진기관과 유방촬영장비가 급격히 증가했기 때문이다. 유방암 검진기관은 2007~2011년에 1374개에서 2042개, 유방촬영장비는 2005~2011년에 1630대에서 2731대로 증가(1, 2)한 반면, 영상 의학과 의사는 2007~2011년에 약 2700명에서 3200명으로 증가하였으며 유방암검진기관이나 유방촬영장비에 비해 증가 폭이 작았다(3). 유방암검진 수검률도 2001~2011년에 약 3%에서 약 60%로 급격하게 증가하였으며 5대암 중 수검률 증가 폭이 가장 크다(1, 4, 5). 이를 통해 영상 의학과 의

사의 판독업무 중 유방촬영의 비중이 최근 10년 동안 급격히 증가한 것을 알 수 있다. 그러므로 지금은 유방영상을 전공하지 않더라도 영상 의학과 의사라면 누구나 유방촬영을 제대로 판독할 줄 알아야 하는 시대가 되었으며, 단순흉부촬영이나 복부초음파 검사처럼 유방촬영도 일반 영상 의학과(general radiology)의 주요 분야라고 할 수 있다.

국가암검진의 암발견율은 0.03~0.12%에 불과하고 그 중 유방암 발견율이 가장 낮다(4, 5). 국가암검진의 유방암 민감도는 약 35%이며 이는 5대암 중 가장 낮은 수치이다(5). 그러므로 유방암검진의 성적을 향상시키기 위해서는 장비에 대한 정도관리와 더불어 판독 질 관리가 필수적이다. 그런데 유방영상 전문의 수는 제한적이기 때문에 전공의 기간 동안 선별유방

촬영 판독에 대한 충분한 수련이 필요하다.

미국영상의학회(American College of Radiology; 이하 ACR)에서는 영상의학과 전공의가 최소 12주 동안 유방영상을 수련받도록 규정하고 있다(6). 그러나 대한영상의학회의 수련 교과과정에는 아직 분야별 수련기간에 대한 명확한 규정이 없는 실정이다(7). 최근 대한영상의학회 수련위원회에서 분야별 최소 수련기간을 설정하고자 준비 중이므로(8) 적절한 수련기간 산정을 위한 객관적인 자료가 필요하다.

저자들이 알기로 우리나라 영상의학과와 유방영상 수련, 특히 유방촬영 판독 수련에 대한 연구는 전무하였다. 이에 저자들은 영상의학과 전공의의 선별유방촬영 판독 성취도를 평가하고 판독 성취도에 영향을 미치는 요인을 분석하였으며 이를 통해 유방영상의 적정 수련조건을 제시하고자 하였다.

## 대상과 방법

58개 영상의학과 수련기관 중 21개 대학병원의 영상의학과 전공의를 대상으로 2010년 6월부터 11월까지 선별유방촬영 판독평가와 설문조사를 시행하였다. 유방촬영 25 증례의 Digital Imaging and Communications in Medicine (DICOM) 파일과 CD viewer (PiViewSTAR™; INFINITT, Seoul, Korea)가 담긴 CD 2장을 각 기관의 유방담당 교수에게 직접 전달하였다. CD와 함께 엑셀파일 형태의 판독평가 답지와 워드파일 형태의 설문지를 같이 전달하였다.

각 기관의 유방영상 담당교수가 유방판독실의 Picture Archiving and Communication System (PACS) 컴퓨터에 DICOM 파일과 CD viewer를 설치하고, 판독평가 답지와 설문지를 출력하여 전공의에게 배부하였다. 판독평가는 전공의가 편한 시간을 택하여 평소 유방촬영 판독환경과 같은 조건에서 판독하도록 하였다. ACR Breast Imaging-Reporting and Data System (BI-RADS)에서 범주(category) 0은 유방촬영만으로는 결론을 내릴 수 없고(inconclusive) 추가 검사나 이전 유방촬영과의 비교가 필요한 경우이다. 선별 유방촬영의 목적은 유방암을 진단하는 것이 아니라, 소환(recall) 여부를 결정하는 것이므로 범주 3, 4, 5를 모두 범주 0에 포함시키는 것이 원칙이다. 그러나 이번 연구에서는 전공의들이 유방암의심 소견을 제대로 인지하고 있는지 평가하기 위하여 범주 4 이상의 유방암의심 증례만 소환하도록 하였다. 양측 유방을 비교하면서 판독하되, 양측 유방의 소환 여부를 개별적으로 표시하도록 하였으며 총 50개의 답을 얻었다. 소환이 필요한 유방은 O, 그렇지 않은 유방은 X로 답지에 표시하였다. 모든 전공의가 판독평가 답지와 설문지 작성을 완료하면 유방영상 담당교수가 답지와 설문지를 취

합하여 연구자에게 우편으로 발송하였다.

유방촬영은 모두 디지털 장비(Senograph DS®; GE, Buc, France)를 이용하여 시행하였다. 25명 중 유방암 환자는 5명이었는데 그 중 한 명은 양성성이었다. 이들은 모두 본원에서 시행한 선별 유방촬영에서 발견된 무증상 여성이었으며 총생검으로 진단된 후 수술로 최종 확진되었다. 나머지 양성 및 정상 증례들은 모두 1년 이상 시행한 추적검사에서 유방암이 발견되지 않았으므로 임상적으로 진음성(true negative)임을 확인하였다.

설문지 항목은 성별, 연차, 소속 수련기관 등의 인적 사항, 유방영상 수련여부 및 수련받은 개월 수(수련기간), 최근 유방영상 수련시기 등의 수련 현황, 그리고 최근 유방영상 수련시기와 판독평가의 간격(수련-평가 간격), 유방영상을 한 달 수련받은 동안 판독하는 유방촬영 건수(판독량), 성취도 평가를 위한 판독에 소요되는 시간(판독시간) 등을 포함하였다. 이와 별도로 유방영상 담당교수를 대상으로 전화설문을 하였는데 해당 기관의 유방영상 총 수련기간, 유방촬영장비의 종류, 유방영상 전임여부와 경력 등을 조사하였다. 유방영상 담당교수가 다른 파트를 겸임하는 경우는 해당 기관의 유방영상 총 수련기간을 1/2로 나누어 산정하였다. 유방영상 경력은 전임의를 포함하여 유방영상에 처음 입문한 시점부터 판독평가 시점까지의 햇수로 산출하였고, 유방영상 담당교수가 여러 명인 경우는 평균값을 구하였다. 그밖에 수련기관 규모는 1000병상을 기준으로 대형기관과 그외 기관으로 나누었고, 수련기관의 위치가 서울을 포함할 수도권인지 지방인지 나누어 분석하였다.

유방촬영 판독평가 답지를 채점하여 진양성(true positive)과 진음성 증례의 합을 판독점수로 정하였다. 수련군과 비수련군으로 나누어 판독 성취도를 비교하였다. 수련군의 정의는 유방영상을 한 달 이상 수련받은 전공의로 정하였고, 한 달 미만으로 수련받았거나 아직 수련받지 않은 전공의는 비수련군으로 분류하였다. 유방영상 담당교수가 다른 파트를 겸임하는 경우는 한 달 수련받은 전공의를 수련군에 포함시키되 유방영상 수련기간은 0.5로 산정하였다. 판독 성취도 평가항목은 판독점수, 소환율(recall rate), 민감도(sensitivity), 양성예측도(positive predictive value), 위양성률(false positive rate)로 정하였다(9). 전공의 전체, 수련군과 비수련군에서 각각 연차별 판독 성취도를 비교하였다. 각 기관의 수련군 전공의들과 비수련군 전공의들의 평균 판독점수의 차이를 수련효과(training effect)로 정의하였다. 각 기관의 수련효과를 측정할 후 수련기관들을 효과군과 비효과군으로 나누었고, 수련군 전공의들을 수련 효과기관(수련 효과군)과 비효과기관에 근무하는 경우(비효과군)로 세분한 후 판독 성취도에 영향을 미치는 인자들을 분석하기 위하여 다음의 변수를 비교하였다: 전공의의 유방영상 수련기간, 한

달 간 판독건수, 수련-평가 간격과 판독시간, 수련기관의 규모, 위치, 유방영상 총 수련기간과 유방촬영장비의 종류, 유방영상 담당교수의 전담여부와 경력 등을 분석하였다. 또한 25 증례 중 가장 흔하게 나타난 진양성, 위음성, 위양성 증례를 분석하였다.

통계학적 분석은 다음과 같이 시행하였다; 판독 성취도 평가 항목별 평균 비교는 unpaired *t*-test를 이용하였고, 수련효과에 영향을 미치는 인자를 분석하기 위하여 모수 변수(parametric data)일 경우 unpaired *t*-test와 Mann-Whitney U test, 비모수 변수(non-parametric data)일 경우 Fisher's exact test와 Pearson chi-square test를 이용하였다. 통계 프로그램은 SPSS 17.0 (SPSS Inc., Chicago, IL, USA)을 이용하였으며 유의 수준은 0.05로 정하였다.

## 결과

연구에 참여한 21개 수련기관 중 20개 기관의 자료를 분석하였다. 한 기관은 전공의의 판독평가 답지가 모두 일치하였기 때문에 부정행위로 간주하여 제외하였다. 총 203명의 전공의가 참여하였으며 이는 영상의학과 전체 전공의의 36.3%에 해당하였다.

연구에 참여한 전공의의 특성은 다음과 같다; 성별 분포는 남자 87명(42.9%)과 여자 116명(57.1%)이었으며, 연차별 분포는 1년차 55명(27.1%), 2년차 53명(26.1%), 3년차 49명(24.1%), 4년차 46명(22.7%)이었다. 수련군과 비수련군은 각각 121명(40.4%)과 82명(40.4%)이었다. 비수련군의 연차별 분포는 1년차 48명, 2년차 25명, 3년차 6명, 4년차 3명이었고, 수련군은 1년차 7명, 2년차 28명, 3년차와 4년차가 각각 43명이었다. 비수련군은 저년차(89.0%)가 많은 반면, 수련군은 고년차(71.0%)가 많았다.

판독평가 당시 수련군은 평균  $1.8 \pm 0.7$ 개월(범위, 0.5~3.0개월)의 수련을 받았는데 그 중 수련기간이 1개월 이하인 전공의는 34명(28.1%), 1.5~2개월 70명(57.9%), 2.5~3개월 17명(14.0%)이었다. 수련군 중 116명이 수련-평가 간격에 대해 응답하였으며 평균  $6.3 \pm 6.7$ 개월(범위, 0~35개월)이었다. 수련군의 한달 평균 판독량은  $564.0 \pm 230.1$ 건(범위, 200~900건)이었었는데 그 중 300건 이하인 경우가 29명(24.0%), 600건 이하 43명(35.5%), 600건을 넘는 경우가 49명(40.5%)이었다. 성취도 평가를 위한 판독에 소요된 시간은 평균  $29.5 \pm 21.4$ 분(범위, 5~270분)이었다. 수련군은 109명이 응답하였고 평균  $28.0 \pm 10.5$ 분(범위, 10~60분)이었으며, 비수련군은 73명이 응답하였고 평균  $31.8 \pm 31.2$ 분(범위, 5~270분)이었다.

연구에 참여한 수련기관의 특성은 다음과 같다; 대형기관이 3개(15.0%), 그 외 기관은 17개였으며, 수도권에 위치한 기관은 8개(40.0%), 지방은 12개였다. 유방영상의 총 수련기간은 평균  $2.3 \pm 0.5$ 개월이었으며 그 중 1.5개월인 기관은 1개, 2개월 11개, 2.5개월 2개, 3개월 6개였다. 유방촬영장비는 디지털(digital) 장비를 사용하는 기관이 18개(90.0%)였고, 나머지 2개 기관은 필름-스크린 장비를 사용하였다. 유방촬영 건수는 한 달 평균  $522.5 \pm 220.9$ 건(범위, 200~900건)이었으며 그 중 300건 이하인 기관은 5개, 600건 이하는 9개였고, 600건을 넘는 기관은 6개였다. 유방영상 담당교수가 유방영상 전임인 기관은 8개였고, 나머지 12개 기관에서는 다른 분야를 겸임하고 있었다. 겸임분야는 갑상선영상 5개 기관으로 가장 많았고, 심혈관영상과 소아영상이 각각 2개, 흉부, 비뇨생식기, 기타가 1개 기관씩 있었다. 유방영상 담당교수의 경력은 평균  $9.6 \pm 4.2$ 년(범위, 4~21년)이었다. 대형기관에 근무하는 경우는 평균  $10.7 \pm 2.5$ 년(범위, 8~11년)이었고, 그 외 기관은 평균  $9.4 \pm 4.5$ 년(범위, 4~21년)이었다. 수도권과 지방 수련기관에 근무하는 유방영상 담당교수의 경력은 각각  $9.4 \pm 1.8$ 년,  $9.7 \pm 5.4$ 년이었다.

25개 증례에서 양측 유방의 소환 여부를 개별적으로 표시하였으므로 선별유방촬영 판독점수는 50점 만점이었다. 전공의 평균은  $40.7 \pm 6.1$ 점(범위, 23~50점)이었고, 수련기관 평균은  $40.7 \pm 3.1$ 점(범위, 35.8~46.3점)이었다. 판독점수, 소환율, 민감도, 양성예측도, 위양성률 등의 판독 성취도 평가항목을 수련군과 비수련군에서 비교한 결과, 수련군은 판독점수(43.1점 vs. 37.1점), 소환율(11.0% vs. 15.5%), 민감도(83.6% vs. 72.0%), 양성예측도(53.0% vs. 32.4%), 위양성률(13.5% vs. 25.2%) 등 모든 항목에서 비수련군보다 성적이 좋았다( $p = 0.000$ )(Table 1). 그러나 수련군과 비수련군 모두 남녀의 차이는 없었다( $p > 0.005$ ). 성취도 평가 항목을 연차별로 비교했을 때 연차가 올라갈수록 모든 항목이 향상되었다( $p = 0.000$ )(Table 2). 그 중 연차와 가장 유의한 상관관계를 보인 항목은 양성예측도( $\rho = 0.414$ )와 판독점수( $\rho = 0.407$ )였다. 그러나 수련군과 비수련군으로 분리하여 연차별 성취도를 비교한 결과 수련군에서는 연차와 민감도가 상관관계( $p = 0.045$ ,  $\rho = 0.183$ )를 보인 반면, 비수련군에서는 연차와 유의한 상관관계를 보이는 항목이 하나도 없었다(Table 2). 수련군을 수련기간에 따라 세분하여 판독 성취도를 비교하였을 때 판독점수, 민감도, 양성예측도는 양의 상관관계를 보였으며, 소환율과 위양성률은 음의 상관관계를 보여, 수련기간이 증가할수록 모든 항목이 향상됨을 알 수 있었다( $p = 0.000$ )(Table 3). 그 중 수련기간과 가장 높은 상관관계를 보인 것은 양성예측도

(rho = 0.504)와 판독점수였다(rho = 0.486).

각 기관에서 수련군 전공의들의 판독점수 평균과 비수련군 전공의들의 판독점수 평균의 차이를 구하여 수련효과를 측정

하였다. 19개 기관에서 수련효과를 측정할 수 있었고 수련효과는 평균  $6.0 \pm 5.5$ 점(범위,  $-2.3 \sim 16.0$ 점)이었다. 한 기관은 수련군만 있어서 수련효과를 계산할 수 없었다. 대부분의 기관

**Table 1. Comparison of Performance in Interpretation of Screening Mammography between Trainee and Non-Trainee Groups**

	Trainee (n = 121)	Non-Trainee (n = 82)	p Value
Score	43.1 ± 4.7	37.1 ± 6.2	0.000
Recall rate	11.0 ± 4.4%	15.5 ± 6.1%	0.000
Sensitivity	83.6 ± 12.9%	72.0 ± 19.8%	0.000
PPV	53.0 ± 20.9%	32.4 ± 17.0%	0.000
FPR	13.5 ± 10.2%	25.5 ± 13.7%	0.000

Note.—FPR = false positive rate, PPV = positive predictive value, Score = score of mammography interpretation

**Table 2. Comparison of Performance in Interpretation of Screening Mammography among All Residents, Trainee and Non-Trainee Groups According to Year**

		All	Trainee	Non-Trainee
Score	1st	37.2 ± 6.0	40.7 ± 4.5	36.7 ± 6.1
	2nd	40.0 ± 6.6	43.3 ± 5.0	36.2 ± 6.3
	3rd	42.6 ± 4.6	42.9 ± 4.4	40.2 ± 5.6
	4th	43.5 ± 4.7	43.5 ± 4.7	44.3 ± 3.8
		p = 0.000, p = 0.407	p = 0.320	p = 0.228
RR	1st	15.1 ± 5.8%	12.1 ± 5.0%	15.6 ± 5.9%
	2nd	13.8 ± 6.2%	10.7 ± 4.4%	17.3 ± 6.2%
	3rd	11.0 ± 4.3%	11.0 ± 4.3%	11.5 ± 5.0%
	4th	10.7 ± 4.5%	10.9 ± 4.6%	8.3 ± 0.6%
		p = 0.000, p = -0.311	p = 0.858	p = 0.363
Sensitivity	1st	69.7 ± 20.3%	73.8 ± 16.3%	69.1 ± 20.9%
	2nd	81.4 ± 12.5%	83.3 ± 12.8%	79.3 ± 21.1%
	3rd	79.9 ± 15.6%	82.1 ± 12.8%	63.9 ± 24.5%
	4th	85.9 ± 14.0%	86.8 ± 11.8%	72.2 ± 34.7%
		p = 0.000, p = 0.312	p = 0.045, p = 0.183	p = 0.218
PPV	1st	32.1 ± 16.7%	39.7 ± 13.3%	30.9 ± 17.0%
	2nd	42.8 ± 19.4%	52.8 ± 18.1%	31.6 ± 13.9%
	3rd	50.7 ± 22.4%	52.5 ± 22.0%	38.5 ± 23.2%
	4th	55.6 ± 21.9%	55.9 ± 22.1%	51.4 ± 22.9%
		p = 0.000, p = 0.414	p = 0.263	p = 0.285
FPR	1st	24.9 ± 13.1%	17.5 ± 10.6%	25.9 ± 13.2%
	2nd	20.3 ± 14.5%	13.0 ± 10.5%	28.5 ± 14.1%
	3rd	14.2 ± 9.9%	13.7 ± 9.7%	17.4 ± 11.6%
	4th	12.7 ± 10.3%	13.0 ± 10.6%	9.1 ± 3.9%
		p = 0.000, p = -0.372	p = 0.512	p = 0.271

Note.—FPR = false positive rate, PPV = positive predictive value, RR = recall rate, Score = mammography interpretation score

**Table 3. Comparison of Radiology Residents' Performance in Interpretation of Screening Mammography According to Training Period**

	0 Month (n = 82)	≤ 1 Month (n = 34)	≤ 2 Month (n = 70)	≤ 3 Month (n = 17)	p-Value
Score	37.1 ± 6.2	41.6 ± 5.6	43.6 ± 4.1	44.0 ± 4.5	0.000
RR	15.5 ± 6.1%	11.9 ± 5.1%	10.6 ± 4.0%	10.4 ± 4.5%	0.000
Sensitivity	72.0 ± 19.8%	79.4 ± 13.6%	85.0 ± 12.4%	86.3 ± 12.1%	0.000
PPV	32.4 ± 17.0%	46.9 ± 20.3%	54.6 ± 20.0%	59.0 ± 23.7%	0.000
FPR	25.5 ± 13.7%	16.4 ± 12.1%	12.6 ± 9.0%	11.8 ± 10.1%	0.000

Note.—FPR = false positive rate, PPV = positive predictive value, RR = recall rate, Score = mammography interpretation score

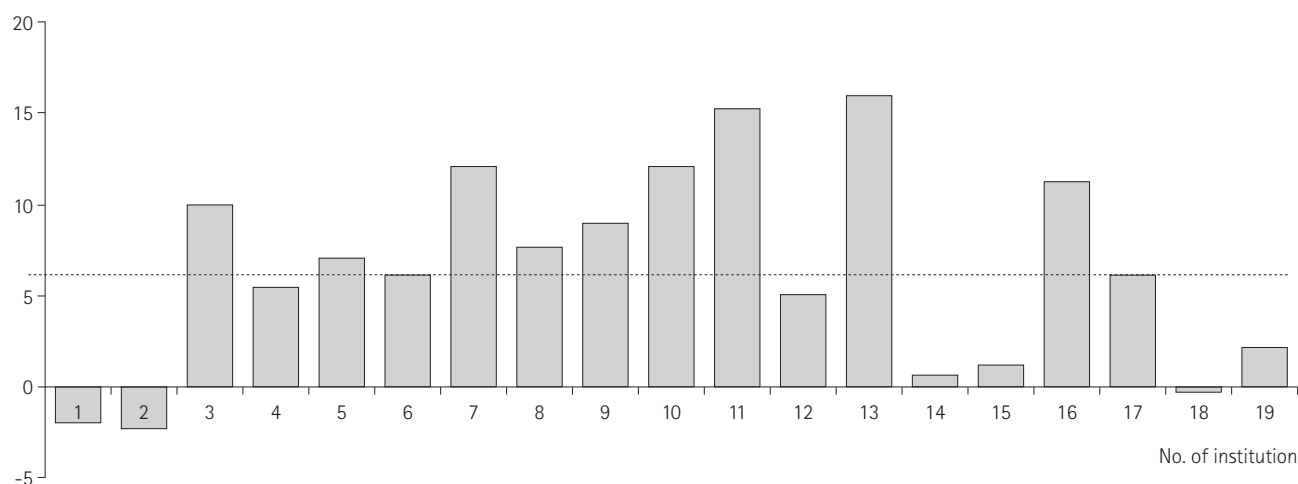


Fig. 1. Radiology residents' training effect of breast imaging according to institutions.

Note.—Dotted line = mean level of training effect, Training effect = score difference between trainee and non-trainee groups in the institution

Table 4. Comparison of Factors Affecting Performance in Interpretation of Screening Mammography between Training-Effective and Non-Effective Groups

		Training-Effective	Non-Effective	p Value
Resident	Training period (mo)	1.8 ± 0.7	1.6 ± 0.7	0.161
	Volume of MG	594.0 ± 242.2	476.9 ± 172.8	0.008*
	Interval (mo)	6.2 ± 6.2	7.6 ± 8.6	0.374
	Test time (min)	27.5 ± 11.1	29.5 ± 8.9	0.407
Hospital	Size	23.1% big	0% big	0.517
	Location	38.5% metro	50.0% metro	1.000
	Total training (mo)	2.2 ± 0.5	2.4 ± 0.5	0.467
	MG type	84.6% DM	100% DM	1.000
Staff	Dedication	61.5%	0%	0.018*
	Experience	8.5 ± 2.1	12.7 ± 6.1	0.087

Note.—\*Statistically significant.

DM = digital mammography, Interval = time interval between breast training and interpretation test, metro = metropolitan area, min = minute, mo = month, Test time = time required to perform interpretation test, Total training = total period of breast training in the institution, Volume of MG = numbers of mammography which a resident reads during one month of training period

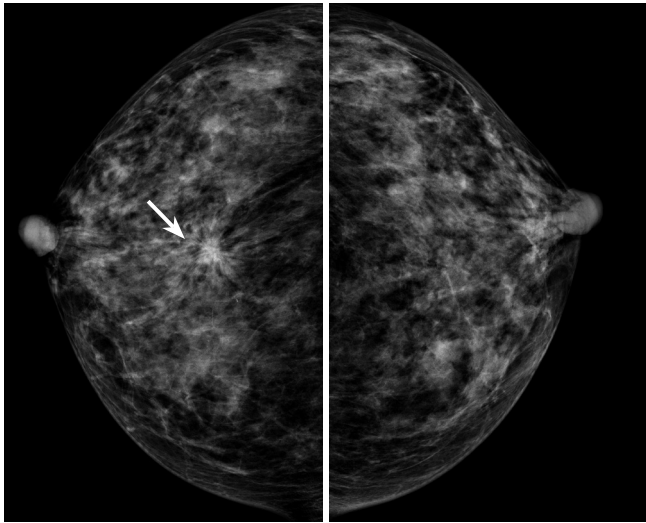
에서 수련효과를 보였으나 그 폭은 5.5~16.0점으로 다양하게 나타났다(Fig. 1). 또한 2개 기관은 수련군 평균이 비수련군 평균보다 오히려 2.0~2.3점 낮았고, 4개 기관은 수련효과가 미미하였다(-0.3~2.2점). 수련 효과군의 수련효과는 평균  $9.5 \pm 3.6$ 점(범위, 5.1~16.0점)인 반면, 비효과군은 평균  $-0.1 \pm 1.8$ 점(범위, -2.0~2.2점)이었다.

수련군 전공의들을 수련 효과기관에 근무하는 경우(수련 효과군)와 비효과기관에 근무하는 경우(수련 비효과군)로 세분하여 판독 성취도에 영향을 미치는 인자가 있는지 분석하였다(Table 4); 수련 효과기관은 비효과기관보다 전공의의 유방촬영 판독량이 많았고(594.0 vs. 476.9건)( $p = 0.008$ ), 유방영상 담당교수가 유방영상만 전임하는 경우가 많았다(61.5 vs. 0%)( $p = 0.018$ ). 수련 효과기관에서는 유방영상 담당교수의

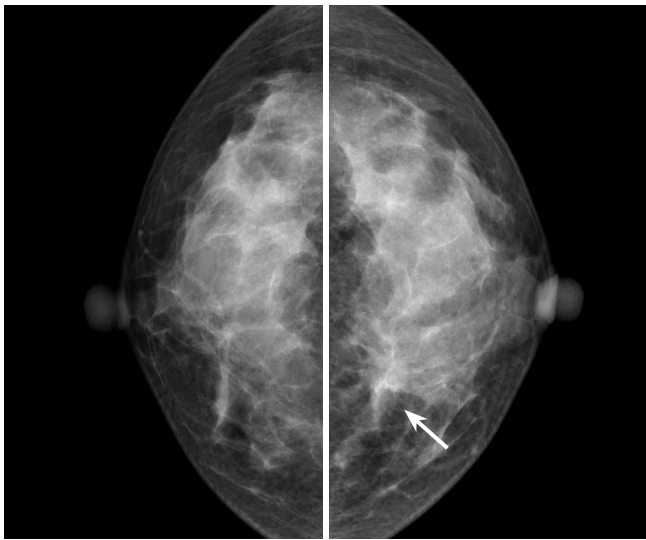
겸임 분야가 갑상선영상이 3개 기관(60.0%), 비뇨생식기와 소아영상이 1개 기관씩이었고, 비효과기관에서는 갑상선영상 2개 기관(33.3%), 흉부, 소아, 심혈관영상과 기타분야가 1개 기관씩 있었다. 전공의의 수련-평가 간격(6.2 vs. 7.6개월), 판독시간(27.5 vs. 29.5분)은 효과군과 비효과군 간에 차이가 없었다( $p > 0.05$ ). 수련기관의 규모(대형 기관, 23.1% vs. 0%)와 지역(수도권, 38.5 vs. 50.0%), 유방영상 총 수련기간(2.2 vs. 2.4개월), 유방영상 담당교수의 경력(8.5 vs. 12.7년)도 양 군 간에 차이가 없었다( $p > 0.05$ ).

증례 중 유방암이 5명 있었는데 한 명은 양측성 유방암 환자였으므로 총 병변 수는 6개였다. 그 중 침윤성 관상피암(invasive ductal cancer)이 5개, 침윤성 소엽암(invasive lobular cancer)이 1개 있었다. 유방암 유형은 종괴 2개, 미세석회화 2개, 미세





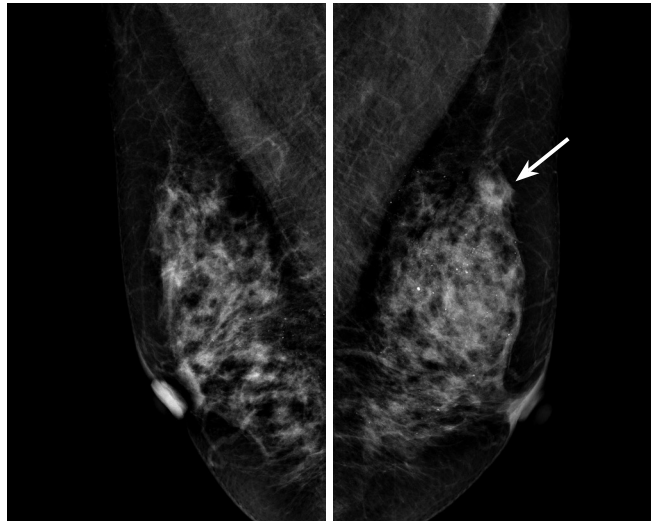
**Fig. 2.** A 56-year-old woman with screening-detected invasive ductal carcinoma in her right breast. Right craniocaudal (CC) view **(A)** shows an irregular spiculated mass (arrow) in the mid portion that is a small but conspicuous cancer. It was the most common true-positive case. However, not all residents interpreted this mass correctly: 95.9% of trainees and 93.9% of non-trainees were correct. Left CC view **(B)** shows negative finding.



**Fig. 3.** A 46-year-old woman with screening-detected invasive lobular carcinoma in her left breast. Right craniocaudal (CC) view **(A)** shows negative finding. Left CC view **(B)** shows an architectural distortion (arrow) with asymmetry in the inner posterior portion. It was the most common false-negative case: 67.8% of trainees and 79.3% of non-trainees missed it.

석회화를 동반한 종괴와 비대칭을 동반한 구조왜곡이 각각 1개 있었다. 유방실질 패턴은 지방형유방이 2명, 치밀유방 3명이었다.

가장 흔한 진양성 증례는 침윤성 종괴였다(Fig. 2). 수련군의



**Fig. 4.** A 56-year-old healthy woman. Right mediolateral oblique (MLO) view **(A)** shows negative finding. Left MLO view **(B)** shows an asymmetric-looking normal breast tissue (arrow) in the upper mid portion. Additionally, diffusely scattered calcifications were evident in both breasts. It was the most common false-positive case: 46.3% of trainees and 40.2% of non-trainees recalled the left breast.

95.9%, 비수련군의 93.9%가 양성으로 판독하였으며 양 군 간에 차이는 없었다( $p = 0.529$ ). 가장 흔한 위음성 증례는 구조왜곡이었다(Fig. 3). 수련군의 67.8%, 비수련군의 79.3%가 음성으로 판독하였고 양 군 간에 차이는 없었다( $p = 0.080$ ). 가장 흔한 위양성 증례는 비대칭 유사음영(asymmetry-like density)이었다. 수련군의 46.3%, 비수련군의 40.2%가 양성으로 판독하였으며 역시 양 군 간에 차이는 없었다( $p = 0.471$ )(Fig. 4).

한편, 유방암 증례 중 수련군과 비수련군 간에 가장 큰 차이를 보인 것은 미세석회화를 동반한 종괴였는데 수련군 중 92.6%, 비수련군 중 72.0%가 양성으로 판독하였다. 반면에 유방암 중 수련군과 비수련군 간에 가장 차이가 없었던 증례는 비대칭을 동반한 구조왜곡이었는데 수련군 중 32.2%, 비수련군 중 20.7%가 양성으로 판독하였다(Fig. 3).

## 고찰

유방암 발견율을 향상시키기 위해서는 유방촬영장비에 대한 질 관리와 함께 영상의학과 의사에 대한 질 관리가 필수적이다. 이를 위해 전문의 대상의 판독 질 관리 교육이나 판독심화 캠프가 필요함과 동시에 전공의 수련기간 동안 유방촬영 판독법을 체계적으로 가르쳐야 한다.

본 연구에서 수련군이 비수련군보다 모든 항목에서 판독 성취도가 높았고, 연차가 올라감에 따라 모든 항목에서 판독 성

취도가 향상되었다. 그런데 본 연구에서 비수련군에는 저년차가 많은 반면, 수련군에는 고년차가 많았기 때문에 판독 성취도 향상이 수련 자체의 효과가 아니라 단순히 연차 효과를 나타낸 것일 수도 있다. 그런 가능성을 배제하기 위하여 수련군과 비수련군 비율이 비슷한 2년차를 비교해 보았을 때, 모든 항목에서 수련군이 비수련군보다 우수함을 확인할 수 있었다. 또한, 수련군 2년차는 비수련군 3년차보다 모든 항목에서 성적이 좋았다. 그러므로 수련군의 판독 성취도는 단순한 연차 효과가 아니라 유방영상에 대한 수련효과 자체를 반영한다는 것을 알 수 있다.

판독 성취도 항목과 연차별 상관관계를 수련군과 비수련군으로 분리하여 살펴보면 유일하게 수련군의 민감도만 연차와 유의한 상관관계가 있었다. 그밖에 양성예측도도 통계적으로 유의하지는 않았지만 모든 연차에서 수련군이 비수련군보다 우수하게 나타났다. 이는 수련군이 비수련군보다 유방암 인지 능력이 우수함을 의미한다. 반면에 4년차는 비수련군이 수련군보다 판독점수, 소환율, 위양성률이 더 좋았는데 이것은 비수련군 4년차가 수련군 4년차보다 더 우수하다는 뜻은 아니다. 이것은 수련군 전공의들이 현재 완성된 상태가 아니라 아직 학습 시기(learning period)에 있기 때문에 위양성을 배제하는 능력이 불완전함을 의미하는 것이다. 즉, 비수련군은 유방암에 대한 지식 자체가 부족하기 때문에 소환율과 위양성률이 낮은데다, 본 연구에서 유방암 증례수보다 정상 증례수가 더 많았기 때문에 판독점수가 오히려 더 높게 나타난 것이다. 본 연구에서 수련군은 평균 1.8개월 수련받고 있는 것으로 나타났는데 4년차 비교 결과, 1.8개월 동안 유방암 인지 능력은 획득할 수 있지만 위양성을 배제하는 능력을 쌓기에는 수련기간이 부족하다는 것을 부수적으로 알 수 있다.

한편, 전공의들 간에 판독점수의 편차가 컸으며( $40.7 \pm 6.1$  점; 범위, 23~50점), 수련기관들도 판독점수의 편차가 큰 것으로 나타났다( $40.7 \pm 3.1$  점; 범위, 35.8~46.3점). 특히, 수련기관 중 31.6%는 수련효과가 아예 없거나 미미하게 나타났기 때문에 유방영상 수련에 대한 개선이 필요함을 알 수 있었다. 수련 효과군에서도 역시 수련효과의 편차가 크게 나타났으므로( $9.5 \pm 3.6$  점; 범위 5.1~16.0점) 유방촬영 판독 수련에 대한 표준화가 필요하다고 생각한다.

본 연구에서 수련효과와 관련이 있는 요인은 전공의의 유방촬영 판독량과 유방영상 담당교수의 전임여부인 것으로 나타났다. 즉, 유방영상만 전임하는 교수의 지도 아래 한 달 평균 약 600건의 유방촬영을 판독하는 전공의는 그렇지 않은 경우보다 판독성적이 좋았다. 반면에 수련기간, 수련-평가 간격, 판독시간, 수련기관의 규모와 위치, 총 수련기간, 담당교수의 경력 등

은 효과군과 비효과군 간에 의미있는 차이가 없었다.

Miglioretti 등(9)은 영상의학과 의사의 선별유방촬영 판독 학습곡선(learning curve)을 보고하였다; 유방영상 전임의를 하지 않은 경우는 영상의학과 의사로서 일하는 첫 해의 소환율이 평균 14%, 위양성률 14%, 민감도는 85%인 반면, 전임의 과정을 마친 경우는 민감도는 같았지만 소환율과 위양성률은 평균 10%로 더 낮았다. 또한 전임의를 하지 않은 경우는 첫 3년 동안 성취도가 점차 증가하여 약 3년의 학습기간이 필요하였으나 전임의 과정을 마친 경우는 별도의 학습기간이 필요 없었다. 본 연구에서 수련군의 소환율은 평균 11.0%, 위양성률은 13.5%였으며 Miglioretti 등이 보고했던, 일반 영상의학과 의사의 첫 해 성적과 비슷하게 나타났다. 또한 수련군의 민감도는 평균 83.6%로 Miglioretti 등(9)의 보고와 비슷하였고 ACR BI-RADS (10)의 권고치에도 근접하므로 수치만 가지고 비교한다면 우리나라 영상의학과 전공의는 대체로 유방영상을 적절하게 수련받고 있다고 생각할 수 있다.

본 연구에서 수련기간이 증가할수록 위양성률이 의미있게 감소하였다(1개월 16.4%, 2개월 2.6%, 3개월 11.8%;  $\rho = -0.446$ ). 그러나 미국 영상의학과 의사들의 선별유방촬영 위양성률이 약 7~8%인 점을 감안할 때(9) 그에 미치지 못하는 수준이다. 높은 위양성률은 국민의료비 증가뿐만 아니라 유방암검진의 질 저하와 직결되므로 전공의의 유방촬영 판독에 대한 수련을 좀더 강화할 필요가 있다.

본 연구에서 수련기간이 증가할수록 판독 성취도가 높게 나타났다지만 4년이라는 한정된 기간 내에서 유방영상 수련기간을 무작정 늘릴 수는 없는 실정이다. Bassett 등(11, 12)에 의하면 2003년에는 미국 영상의학과 수련기관 중 93%가 유방영상을 8주 이상 수련시켰고, 73%는 12주 이상 수련시켰지만, 2011년에는 12개월 이상 수련시키는 기관이 95%로 증가하였고 19%의 기관에서는 유방영상을 18주 이상 수련시키고 있다. 그런데 본 연구에 포함된 수련기관 중 유방영상 수련기간이 3개월인 기관은 25%에 불과하였으므로 총 수련기간이 3개월 미만인 기관은 3개월로 상향조정이 필요하다.

본 연구에서 가장 흔하게 나타난 진양성 증례는 침윤성 종괴였다. 치밀유방이고 병변의 크기가 작았지만 전형적인 유방암의 형태이므로 영상의학과 의사라면 절대 간과해서는 안 되는 병변이다. 그러나 수련군의 4%는 이 병변을 발견하지 못하였다. 가장 흔한 위음성 증례는 비대칭을 동반한 구조왜곡이었는데 수련군의 68%가 병변을 인지하지 못했다. 이는 우리 전공의들이 구조왜곡에 약하다는 것을 의미한다. 가장 흔한 위양성 증례는 비대칭 유사음영이었는데 수련군의 46%가 유방암의심 병변으로 인지하였다. 이는 우리 전공의들이 의미있는 비대칭

음영과 비대칭 유사음영을 제대로 구분하지 못함을 의미한다. 그러므로 유방촬영 판독에 대한 경험을 늘리고 수련을 강화할 필요가 있다.

본 연구의 제한점은 첫째, 전체 58개 영상의학과 수련기관 중 20개 대학병원만 연구에 참여하였으므로 유방영상 수련환경이 부실할 가능성이 높은 기관들은 대부분 포함되지 않았다는 점이다. 그러나 수련기관의 지역적 분포와 각 대학병원의 규모를 감안하여 적절하게 안배하였으며, 전체 전공의의 1/3 이상이 본 연구에 참여하였기 때문에 충분히 대표성이 있는 연구라고 생각한다. 둘째, 판독 성취도를 평가하기 위한 증례수가 25개에 불과하므로 본 연구에서 나타난 판독 성취도가 전공의의 실제 실력을 제대로 반영하는 데 한계가 있을 수 있다는 점이다. 그러나 전공의들의 피곤한 현실을 감안할 때 증례수가 많아질수록 참여도나 집중도가 현저히 떨어질 가능성이 있다. 그러므로 25문제는 아주 충분한 양은 아니지만 일회성 평가로는 적절한 양이라고 생각한다. 실제로 최근 수년간 시행된 유방촬영 판독 질 관리를 위한 교육이나 유방촬영 판독심화캠프에서도 교육 효과를 평가하기 위해 25문제를 이용하고 있다. 셋째, 소환 여부를 결정할 때 일반적으로 좌우 구분을 하지 않는 것과 달리 본 연구에서는 좌우 유방의 소환 여부를 개별적으로 표시했다는 점이다. 이것은 실제 판독환경과는 다르지만, 우연히 맞추는 결과를 피하고 평가의 정확도를 높이기 위함이었다. 예를 들어, 전공의가 한쪽 유방암을 인지하지 못하고 대신 반대편의 의미 없는 소견을 소환한다면 소환 여부는 맞췄지만 실제로는 틀린 것이므로 이런 경우를 평가에 정확하게 반영하기 위해서였다. 마지막으로, 연구자 한 명이 모든 증례를 선정했기 때문에 증례 선정 과정에 선택 오류(selection bias)가 있을 수 있고, 문제의 난이도가 적절하지 않았을 가능성이 있다. 이런 문제를 피하기 위해서는 사전 검증과정이 반드시 필요한데 경험 부족으로 인해 연구계획 수립 단계에서 미처 고려하지 못하였다.

본 연구는 영상의학과 전공의의 선별유방촬영 판독 수련과 관련한 국내 최초의 연구이며, 유방암검진의 확대에 따른 영상 의학과 의사의 역할 변화라는 근래의 실제 진료현장을 반영하였다. 또한 국내 영상의학과 연구로는 매우 드물게 전국 규모의 다기관 연구이므로 위의 몇 가지 제한점에도 불구하고 충분히 가치있는 연구라고 생각한다.

결론적으로, 수련군이 비수련군보다 판독 성취도가 높았고, 수련기간이 길수록 판독 성취도가 높았다. 또한 연차가 올라갈수록 성취도가 높았으나 이는 수련군에만 해당하였다. 수련효과라는 평균 6점이었으나 전공의별, 수련기관별 편차가 크게 나타났다. 수련효과에 영향을 미치는 인자는 전공의의 유방촬영 판독량과 유방영상 담당교수의 전임여부였다. 선별유방촬영 판

독을 위한 바람직한 수련조건은 유방을 전임하는 교수의 지도 하에 월평균 약 600건의 유방촬영을 판독하는 것이며, 수련기간은 2개월보다는 3개월이 바람직하다.

## 감사의 글

데이터 수집을 도와주신 여러 교수님들께 감사드립니다.

가톨릭대학교 서울성모병원 강봉주, 건양대학교병원 김금원, 경북대학교 칠곡병원 김혜정, 계명대학교 동산의료원 우성구, 고려대학교 안암병원 조규란, 단국대학교병원 김유미, 동국대학교 일산병원 김대봉, 동아대학교병원 이진화, 부산대학교 양산병원 추기석, 성균관대학교 삼성서울병원 한부경, 순천향대학교 서울병원 장운우, 순천향대학교 천안병원 김형환, 영남대학교병원 황미수, 울산대학교 서울아산병원 차주희, 원광대학교병원 김혜원, 을지대학교 대전을지병원 김규순, 인제대학교 부산백병원 박영미, 전남대학교 화순병원 임효순, 차의대 분당병원 고경희.

## 참고문헌

1. Park EC. Extending an early cancer screening program: a study for construction of infra-structure and quality control. *National cancer center*, 2005
2. Paik SH. A study for correlation of assessment of special medical equipments and compensation of medical cost. *Health insurance review and assessment service*, 2012
3. Kim BS. Introduction of resident training system in Korean society of radiology. *Orientation for new radiology residents*, 2011
4. Shin HR. Assessment of outcome of early cancer screening program and setup of quality control system. *National cancer center*, 2002
5. Jun JK. The past, present and future of national cancer screening program. Workshop for quality control in mammography interpretation for national cancer screening program, 2011
6. Mammography Quality Standards Act (MQSA). Available at: <http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActProgram>. Accessed 02 March 2011
7. Year-specific details for residents training and teaching program. A notice from Ministry of Health 2011-174 \_Re-



- vised\_2011.12.29
8. Resident training program. Workshop for members of committee of resident training in Korean society of radiology, 2010
  9. Miglioretti DL, Gard CC, Carney PA, Onega TL, Buist DS, Sickles EA, et al. When radiologists perform best: the learning curve in screening mammogram interpretation. *Radiology* 2009;253:632-640
  10. D'Orsi CJ, Bassett LW, Berg WA, Feig SA, Jackson VP, Kopan DB, et al. BI-RADS: mammography, 4th ed. In D'Orsi CJ, Mendelson EB, Ikeda DM. *Breast Imaging Reporting and Data System: ACR-BIRADS- breast imaging atlas*. Reston, VA: American College of Radiology, 2003
  11. Bassett LW, Monsees BS, Smith RA, Wang L, Hooshi P, Faria DM, et al. Survey of radiology residents: breast imaging training and attitudes. *Radiology* 2003;227:862-869
  12. Bassett LW, Bent C, Sayre JW, Marzan R, Verma A, Porter C. Breast imaging training and attitudes: update survey of senior radiology residents. *AJR Am J Roentgenol* 2011;197:263-269

## 영상의학과 전공의들의 선별유방촬영 판독 성취도<sup>1</sup>

이은혜<sup>1</sup> · 류채연<sup>2</sup>

**목적:** 영상의학과 전공의들의 선별유방촬영 판독 성취도를 평가하고 수련효과 관련인자를 분석하였다.

**대상과 방법:** 21개 영상의학과 수련기관의 전공의 203명을 대상으로 선별유방촬영 판독평가와 설문조사를 시행하였다. 수련군과 비수련군에서 판독점수, 소환율, 민감도, 양성예측도, 위양성률 등 판독 성취도를 비교하였고, 각 기관의 수련군 전공의들과 비수련군 전공의들 간의 평균 판독점수 차로 수련효과를 측정하였다. 수련기관을 수련 효과군과 비효과군으로 나누어 판독 성취도 관련인자를 분석하였다.

**결과:** 수련군이 비수련군보다 판독점수(43.1 vs. 37.1), 소환율(11.0 vs. 15.5%), 민감도(83.6 vs. 72.0%), 양성예측도(53.0 vs. 32.4%), 위양성률(13.5 vs. 25.5)이 높았다. 수련기간이 길수록 판독점수, 민감도, 양성예측도는 올라가고( $\rho = 0.486, 0.343, 0.504$ ), 소환율과 위양성률은 감소하였다( $\rho = -0.375, -0.446$ ). 수련효과는 평균 6점이었으나 31.6%의 기관은 수련효과가 없었다. 수련기관 중 수련 효과군과 비효과군은 전공의의 한 달 간 유방촬영 판독량(594.0 vs. 476.9)과 유방영상 담당교수의 전임여부(61.5 vs. 0%)에 차이가 있었다.

**결론:** 수련군이 비수련군보다 선별유방촬영 판독 성취도가 높았으며, 수련기간에 비례하여 판독 성취도가 향상되었다. 수련효과 관련인자는 유방촬영 판독량과 유방영상 담당교수의 전임여부였다.

<sup>1</sup>순천향대학교 의과대학 부천병원 영상의학과, <sup>2</sup>본당서울대학교병원 영상의학과