

## 교육이 유방 초음파의 전문성에 미치는 영향

연세대학교 의과대학 영상의학과, 방사선의과학연구소, 연세대학교 의과대학 의학통계학과\*

백송이 · 김민정 · 김은경 · 최은희\* · 구혜령 · 김지연 · 김한솔 · 박성호 · 윤정현 · 이미정 · 이승우 · 박진영

### = 초록 =

**목적** : 전문의와 함께 유방초음파의 BI-RADS에 대한 교육을 받기 전과 후를 비교함으로써 유방 초음파에 있어서의 수련의 효과를 알아보하고자 한다.

**대상 및 방법** : 진단의 성적을 평가할 자료는 임상정보 없이 조직학적으로 증명된 100개의 병변 (54: 양성, 46: 악성)의 초음파 사진으로 만들어졌다. 두 명의 전문 평가자는 각자 증례를 판독하고, 일치된 의견을 이루었다. 6명(2년차, 4: 3년차, 2)의 전공의들은 1달 동안 전문의와 함께 유방 초음파를 시행, 판독하고 교육 파일들을 공부 하였다. 교육을 받기 전, 후에 BI-RADS에 따라서 평가 자료를 보고 병변의 중피의 모양, 방향성, 변연, 병변 가장자리, 내부 에코, 후방 에코 소견, 주변 조직, 석회화와 최종 평가 범주를 정하였다. 교육 전과 후의 수련의 성과는 일치도와 진단의 정확도로 알아보았다.

**결과** : 교육을 받고 나서, 2년차 전공의들은 통계적으로 유의하게 36개 항목 중 15개 항목에서 전문의와의 일치도의 증가를 보였으며, 3년차 전공의들은 18개 항목 중 3개의 항목에서 통계적으로 유의한 일치도의 증가를 보였다. 정확도에 있어서 전공의들의 평균 민감도는 교육 전에 94.2% (89.1-100%)이었으나 교육 후, 97.8% (95.7-100%)로 상승하였다. 평균 특이도는 교육 전 50.7% (40.7-59.3%)에서 교육 후 40.0% (31.5-51.9%)로 감소되는 결과를 보였다.

**결론** : 본 연구 결과에 따르면, 유방 초음파에 관한 교육은 전공의들이 병변을 해석하고 최종 평가를 내리는데 있어서 일치도를 향상시켰다.

색인단어 : Breast US; BI-RADS, training

## 서 론

유방 초음파 검사는 유방촬영술의 보조적인 진단 방법뿐만 아니라, 치밀유방에 있어서는 유방암의 선별검사로서 매우 중요한 역할을 하게 되었다 [1]. 미국에선 유방촬영술을 판독하는 의사의 자격을 법적으로 규정하였으며, 반드시 영상의학과

전문의이어야 하고, 유방촬영술을 처음 판독하는 의사는 감독자의 관리하에 6개월 동안 약 240여건의 유방촬영술을 판독한 경험이 있어야 하며, 전공의도 이에 준해야 한다는 규정하고 있다. 유방 초음파검사도 미국 Society of Breast Imaging에서는 전공의와 전임의에 대해서 교육 방침을 정하여 초음파의 원리, 시행 방법, 적응증, 초음파 해부학과 병변 해석의 숙달, 석회화 진단에 있어서의 제한점, 유방 촬영술과 임상 정보와의

관계 등을 배우도록 권고하고 있다 [2, 3]. 유방촬영술에 대해서는 이미 Breast Imaging Reporting and Data System (BI-RADS) 교육이 전공의들이 전문가와 최종 평가에 대해 일치된 의견을 내는데 도움이 된다고 알려져 있다 [4, 5]. 하지만 유방 초음파에서는, 병변을 최종 평가 하는데 있어서 BI-RADS를 배우고, 전문의의 초음파 결과 해석과 본인의 해석을 비교해 보며, 의견 일치를 이루는 경험을 하는 것이 전공의들에게 얼마나 도움이 되는 지에 대해서는 아직 연구된 논문이 없다.

본 연구에서는, 교육 전과 후에 영상의학과 전공의가 유방 초음파 검사에서 발견된 병변에 대해, BI-RADS 어휘를 사용하고 최종 평가 범주를 결정 한 결과가 유방 영상 전문의들의 판독 결과와 비교하여, 일치도와 정확도가 어떻게 변하는지를 알아보고자 한다.

## 대상과 방법

본원 의료윤리 위원회에 조인을 구했으며, 영상과 데이터 베이스에서 모든 환자에 대한 정보는 제거되었으므로, 환자에 대한 사전동의는 면제되었다.

2006년 7월-2006년 12월까지 유방영상학 파트에서 1달씩 수련한 전공의 2년차 4명, 3년차 2명 등 총 6명의 영상의학과 전공의들이 연구의 피험자 (이하 평가자)가 되었다. 2년차 전공의는 이전에 유방 초음파와 유방 촬영에 대한 경험은 없으며, 복부와 갑상선 초음파에 대한 경험이 2달, 그리고 1년 동안 유방 분야에 대한 강의를 15차례 정도 들은 경험이 있었다. 3년차 전공의는 이전에 1달간 유방 초음파와 유방 촬영에 대한 경험이 있으며, 복부와 갑상선 초음파에 대한 경험은 3달, 그리고 2년 동안 25차례 정도의 유방 분야 강의를 들었다. 평가자에 대한 교육은 유방영상학에 10년, 4년 경력이 있는 두 명의 영상의학과 전문의에 의해 지도가 이루어 졌다. 이들은 1년간 최소한 3,000건 이상의 유방촬영술을 판독하며, 2,000건 이상의 유방 초음파 검사를 시행, 판독한 경험이 있었다. 평가자들의 교육은 아래와 같이 기준을 두어 이루어 졌다. 1) 200 여건 이상의 유방 초음파를 실시간 검사 및 판독을 하였고 2) 100 여건 이상의 유방 촬영술을 판독 하였다. 3) 미국 방사선 의학회(American College of Radiology, ACR)에서 발행한 유방초음파검사에 대한 BI-RADS를 학습하였고, 4) 평가자는 약 50개의 유방초음파 사례들을 수집, 평가해, 각각의 사례들에 대해 전문의가 해석한 BI-RADS 어휘와 최종평가범주를 본인의 것과 비교하였다. 5) 그 외에 전문의에 의해 진행되는 유방 영상학에 관한 강의를 2시간 이상 들었으며, 유방영상학에 관련된 교육 파워포인트 파일(ppt file)을 5개 이상 학습하였다. 유방영상학에 관련한 교육 파워포인트 파일은 유방영상학에 관한 전반적인 접근 방식과 유방촬영술과 추가 유방촬영술에 대한 이해, 그리고 유방초음파에 관한 내용을 담고 있었다.

## 평가 자료

평가자에게 주어지는 평가 파일은 한 명의 영상의학과 전문의에 의해 수집되었다. 수집된 100에는 모두 조직검사 혹은 2년 이상의 추적검사로 양성(n=54) 혹은 악성(n=46)으로 확진 되었다. 각각의 사례에 대해 두 개의 대표적인 가로축 세로축 초음파 영상을 티프(TIFF, Tag Image File Format)파일로 300인치당 점수(dpi=dots per inch) 이상의 해상도로 전환시키고, 환자의 임상정보는 평가사진에 포함되지 않도록 하였다. 사례들을 무작위 순서로 파워포인트 파일로 정리하였다. 사례 수집에는 관여하지 않고, 수련을 지휘 감독하였던 두 명의 영상의학과 전문의 (이하 전문평가자)가 BI-RADS 어휘와 최종평가범주에 맞추어 사례들을 독자적으로 평가하였으며, 서로 상이한 경우는 상의하여 의견의 일치를 보았다. 평가하는 BI-RADS 어휘 및 최종평가범주는 ACR에서 발행한 BI-RADS에 따랐으며, 최종 평가범주에 있어 BI-RADS에선 9가지의 범주(4a, b, c를 포함한 0에서 6)로 분류되나 본 연구에서는 주어진 병변에 대한 해석이므로 범주 0, 1, 6은 이 연구에서 제외하였다. 평가자들은 평가 파일을 보고 평가 항목 각각에 적절한 항목을 선택하도록 하였으며, 항목은 9가지로 각각은 아래와 같았다. 종괴의 모양 (불규칙, 원형, 난원형), 종괴의 방향성 (피부와 평행, 피부와 평행하지 않은), 변연 (국한성, 불분명한 경계, 미세소엽형 경계, 침상형 경계, 각진형 경계), 병변 가장자리 (갑자기 끝나는 가장자리면, 에코성 테두리), 종괴의 내부 에코 (무에코, 저에코, 동일에코, 고에코, 복합에코), 후방 에코 소견 (무후방 음향, 후방음향 증가, 후방음영, 결합 양상), 주변 조직 (주변 유관 변화, 주변 구조 왜곡, 부종, 쿠퍼씨 인대의 변화, 피부 비후, 피부의 퇴축/불규칙화), 석회화 (종괴 내 미세 석회화, 종괴 바깥의 미세 석회화, 거대 석회화, 석회화 없음)를 포함하는 8개 항목과 최종 평가 범주(2, 3, 4a, 4b, 4c, 5)였다. 각각의 항목의 결과는 전산화된 매트릭스 정산표(Excel, Microsoft, Redmond, WA)에 기재하였다.

## 평가

평가자들은 유방영상학 파트에 임하는 첫날과 수련을 마치는 한 달의 마지막 날에 각각 두 차례(전, 후)에 걸쳐 평가 파일로 평가 하였다. 모든 평가자들에게 이 연구에 포함된 사례들의 양성 대 악성 비율 및 사례들의 임상적 정보나 조직학적 결과에 대한 정보는 주어지지 않았으며, 다른 평가자들의 평가 기록도 주어지지 않았다. 평가자들은 평가 파일을 보고 평가 항목 각각에 적절한 항목을 선택하도록 하였으며, 평가 항목은 전문평가자의 항목과 동일하며, 역시 각각의 항목의 결과는 전산화된 매트릭스 정산표(Excel, Microsoft, Redmond, WA)에 기재하였다.

**자료분석**

2차례 평가에 걸친 평가자와 전문평가자 간의 평가 결과 변이성, 각 평가자의 전평가 및 후평가 결과의 변이성에 대해 카파 통계를 사용하였다 [6, 7]. 카파 값 ( $\kappa$ -value)은 Fleiss등이 제안한 바와 같이 0.40 미만은 적은 일치 (poor agreement): 0.40-0.75, 양호한 일치 (fair - good agreement): 0.75 초과는 우수한 일치 (excellent agreement)로 보았다. 병변의 평가에 있어서 여러 평가자들과 전문 평가자들 사이의 일치도를 보기 위해 일반화된 카파 값을 사용하였다 [7, 8]. 일반화된 카파 값은 두 명을 초과하는 검사자간의 신뢰도를 의미하는 것으로, 양의 값은 평가 변이가 주로 병변의 차이에서 생기며 이들 병변의 차이에서 생기는 변이보다, 검사자들의 일치도가 높음을 의미한다. 음의 값은 반대로 평가자의 불일치에서 평가 변이가 기인함을 의미한다. 0의 값은 두 가지 변이, 즉 평가자의 일치도와 병변의 차이에 의한 기여도가 동등한 것을 의미한다. 같은 평가자에서 각 항목별 일치도 변화는 McNemer test로 유의한 차이가 있는

지를 보았다. 또한 각각 평가자의 전, 후 평가의 민감도, 특이도, 양성예측도와 음성예측도 등은 카이-스퀘어 방법으로, 정확도는 수신자 판 단 특성 곡 선 (Receiver operator characteristic curve: ROC)의 곡선하 면적 (area under curve: Az value)으로 비교하였다. 통계 처리는 SAS system (SAS Institute Inc., Cary, NC, U.S.A.)를 이용하였다.

**결 과**

**일치도 평가**

각 항목에 대한 평가자들의 해석을 전문평가자의 동의에 의한 결과와 비교하여 얻은, 수련 전후의 일치도는 표 1과 같으며, 6명의 평가자에서 각각 BI-RADS 용어 범주 8개 항목과 최종 평가 범주, 총 54개 항목 18개의 항목에서 수련 전후의 전문평가자와의 평가 일치도가 통계적으로 유의하게 증가하였으며, 특히 전공의 2년차 4명, 1년부터 4년까지 (2nd year 1-

**Table 1** Change of Reviewers' Agreement for Lesion Description and Final Assessment after Training

|            | Shape  |         | Orientation |         | Margin |         | Lesion-boundary |         | Echo-pattern |         | Posterior echo feature |         | Calcification |        | Surrounding tissue |         | Final assessment |        |
|------------|--------|---------|-------------|---------|--------|---------|-----------------|---------|--------------|---------|------------------------|---------|---------------|--------|--------------------|---------|------------------|--------|
|            | Pre    | post    | Pre         | post    | Pre    | post    | Pre             | post    | Pre          | post    | Pre                    | post    | Pre           | post   | Pre                | post    | Pre              | post   |
| 2nd year 1 | 0.2214 | 0.4947* | 0.3097      | 0.7254* | 0.3446 | 0.3633* | 0.082           | 0.4605* | 0.5749       | 0.7127* | 0.4933                 | 0.5615* | 0.5498        | 0.6111 | 0.506              | 0.9509* | 0.5743           | 0.4963 |
| 2nd year 2 | 0.5055 | 0.5708* | 0.721       | 0.8599  | 0.4343 | 0.4237  | 0.3423          | 0.4287  | 0.6946       | 0.7151  | 0.4541                 | 0.5277* | 0.6683        | 0.8059 | 0.4201             | 0.6779* | 0.6073           | 0.6188 |
| 2nd year 3 | 0.4168 | 0.6061* | 0.617       | 0.4221* | 0.5001 | 0.442   | 0.356           | 0.6585* | 0.5294       | 0.5458  | 0.5001                 | 0.5312  | 0.7388        | 0.73   | 0.5369             | 0.4201  | 0.5993           | 0.527  |
| 2nd year 4 | 0.4996 | 0.544   | 0.7665      | 0.8039  | 0.2665 | 0.4127  | 0.2938          | 0.6417* | 0.6922       | 0.5851* | 0.5644                 | 0.6464  | 0.7075        | 0.779  | 0.6779             | 0.6779  | 0.623            | 0.6517 |
| 3rd year 1 | 0.4199 | 0.5014  | 0.7762      | 0.6704  | 0.5362 | 0.5553  | 0.2081          | 0.6708* | 0.6836       | 0.6659  | 0.4578                 | 0.4261  | 0.4757        | 0.4942 | 0.6779             | 0.7481  | 0.6026           | 0.6366 |
| 3rd year 2 | 0.3983 | 0.4071* | 0.7974      | 0.7799  | 0.5281 | 0.5429  | 0.4159          | 0.3854  | 0.5813       | 0.7011  | 0.6069                 | 0.5727  | 0.5054        | 0.5475 | 0.4427             | 0.6531* | 0.5693           | 0.601  |

Pre : before training, Post : after training

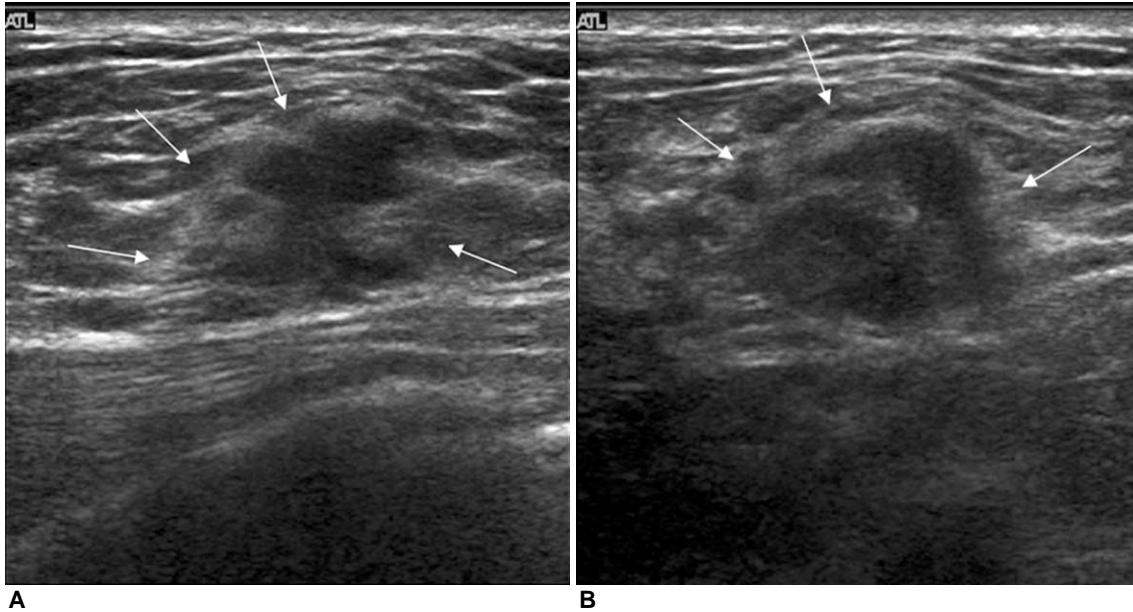
\* : p < 0.05 in McNemer test

**Table 2.** Generalized k-value of before and after Training in Residents

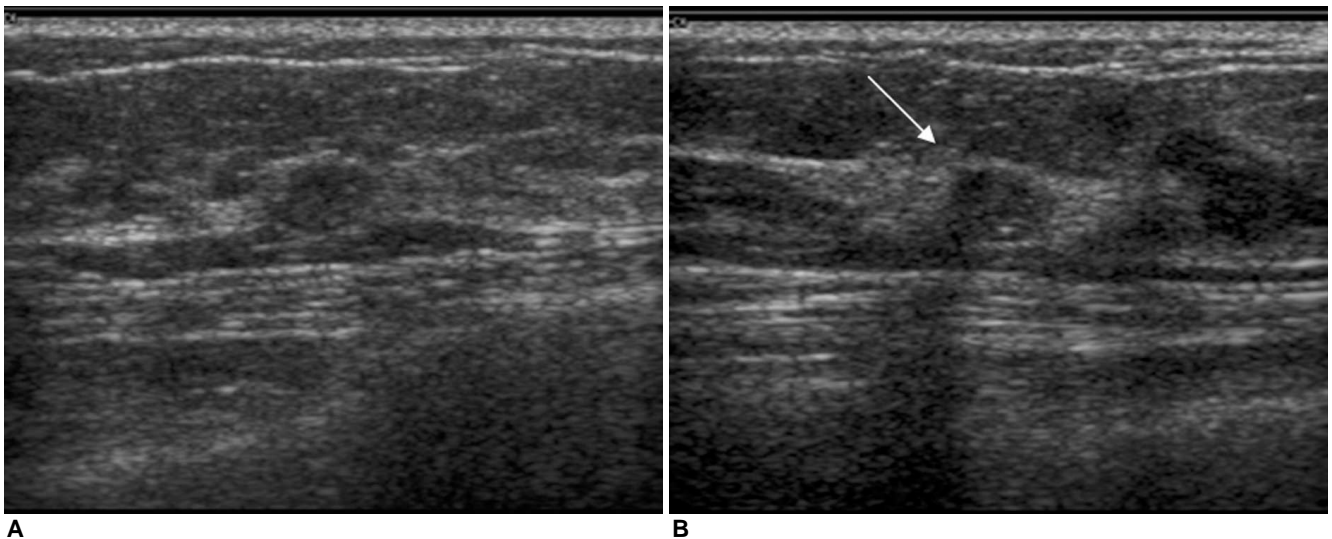
| Variable               | $\kappa$ -value before training | $\kappa$ -value after training | Difference of $\kappa$ - value |
|------------------------|---------------------------------|--------------------------------|--------------------------------|
| Shape                  | 0.44175                         | 0.49801                        | 0.05626                        |
| Orientation            | 0.62662                         | 0.62305                        | -0.00357                       |
| Margin                 | 0.32432                         | 0.31191                        | -0.01241                       |
| Lesion-boundary        | 0.28673                         | 0.47569                        | 0.18896                        |
| Echo-pattern           | 0.48055                         | 0.57593                        | 0.09538                        |
| Posterior echo feature | 0.52762                         | 0.52635                        | -0.00127                       |
| Calcification          | 0.34762                         | 0.39963                        | 0.05201                        |
| Surrounding tissue     | 0.083937                        | 0.15156                        | 0.067623                       |
| Final assessment       | 0.31968                         | 0.33411                        | 0.01443                        |

4)에선 36개의 항목 중 15개의 항목에서 통계적으로 유의한 일치도 증가를 보였다. 전공의들 전체의 수련 전후 일반화된 일치도 변화는 Table 2와 같았다. 평가자들이 수련을 받기 전에 측정된 일반화된 카파 값은 병변 주변 조직 항목에서 가장 낮은 값을 보였고(Table 2), 종괴의 방향성 항목에서 가장 높은 값을 보였다. 한달 간의 수련을 마치고 나서 종괴의 모양,

병변 가장자리, 종괴의 내부 에코, 석회화 그리고 주변 조직 항목에서 카파 값의 증가를 보였다. 특히 병변 가장자리 항목에서 수련 후 카파 값은 0.29에서 0.48로 가장 큰 일치도의 상승 결과를 보였다 ( $p < 0.05$ ) (Fig. 1). 종괴의 방향성과 변연 항목은 수련 후 일치도가 감소하는 결과를 보였으나 통계적인 차이는 없었다. 종괴의 모양, 방향성 그리고 내부 에코는 수련 전



**Fig. 1** Longitudinal (A) and transverse (B) image. Before training, four reviewers thought that this lesion had abrupt interface, but after training three of them changed their minds as this lesion had an echogenic halo (arrowed) at the lesion boundary. Experienced breast radiologists thought that this lesion had an echogenic halo and this lesion was confirmed as malignant. As in this case, the descriptor for lesion boundary was improved after training.



**Fig. 2.** Longitudinal (A) and transverse (B) image. Three reviewers thought this lesion to be a probable benign lesion, category 3, but they changed the category to 4a after training. The remaining 3 reviewers categorized this lesion as 4a, 4b and 4b before training, but they changed their minds as 3, 4a and 4a. Experienced breast radiologists categorized this lesion as category 4a because this lesion had partially spiculated margin (arrowed), and proved to be malignant in pathology. In this case, agreement and accuracy in final assessment was improved after training.

과 후 모두에서 카과 값이 0.4이상으로 양호한 일치도를 보였다. 그러나, 최종 평가 범주에 있어서의 평가자들의 일치도는 수련 전 후 0.32에서 0.33으로 증가하였으나 통계적인 의미는 없었다 (p>0.05).

**정확도 평가**

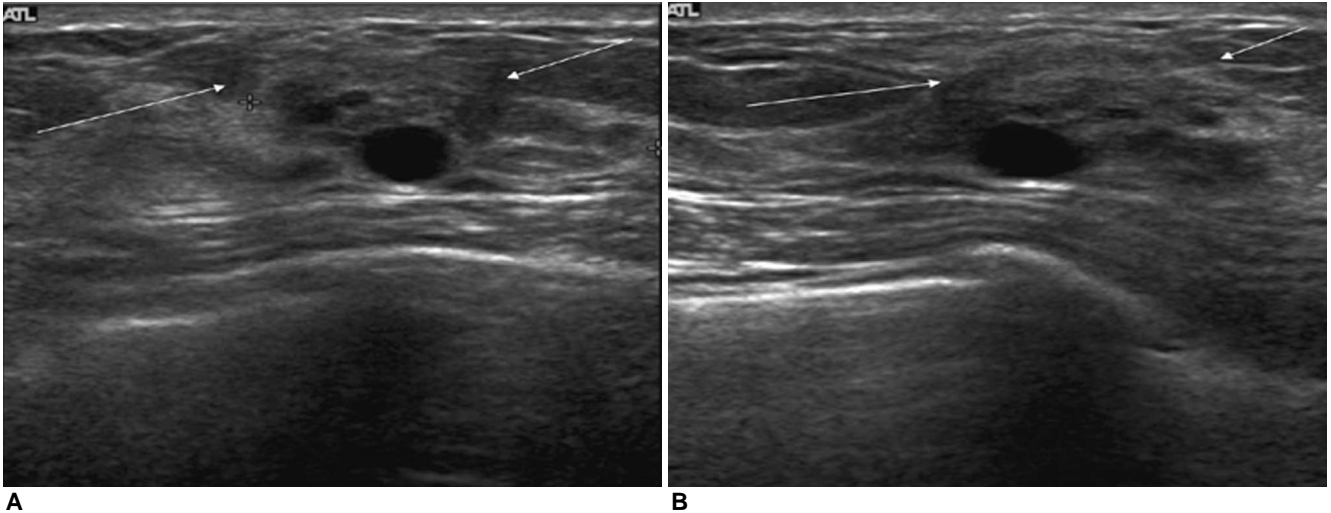
전문 평가자들은 병리학적인 진단을 기준으로 할 때 민감도

97.8%, 특이도 64.8%, 음성예측도 97.2%, 양성 예측도 70.3%의 결과를 보였고, 최종 평가에 있어서 Az value는 0.88이었다. 평가자들의 민감도는 교육 전엔 평균 94.2% (89.1-100%)였으며, 교육 후엔 평균 97.8% (95.7-100%)로 증가하였다 (Table 3, Fig. 2, 3). 6명중 3명에서 증가된 결과를 보였으나 통계적인 차이는 없었다. 특이도는 교육 전에 평균 50.7% (40.7-59.3%)였으며, 교육 후엔 평균 40.0%

**Table 3.** Performance of Residents and Expert according to Histological Reference Standard

|                         |             | Before training (%) | After training (%) |
|-------------------------|-------------|---------------------|--------------------|
| 2nd year 1              | Sensitivity | 89.1                | 97.8               |
|                         | Specificity | 59.3                | 46.3               |
|                         | NPV         | 86.5                | 96.2               |
|                         | PPV         | 0.8320              | 0.8430             |
|                         | Az value    | 65.1                | 60.8               |
| 2nd year 2              | Sensitivity | 95.7                | 95.7               |
|                         | Specificity | 48.2                | 40.7               |
|                         | NPV         | 92.9                | 91.7               |
|                         | PPV         | 61.1                | 57.9               |
|                         | Az value    | 0.8730              | 0.8690             |
| 2nd year 3              | Sensitivity | 95.7                | 100                |
|                         | Specificity | 48.2                | 24.1               |
|                         | NPV         | 92.9                | 100                |
|                         | PPV         | 61.1                | 52.9               |
|                         | Az value    | 0.8700              | 0.8720             |
| 2nd year 4              | Sensitivity | 100                 | 100                |
|                         | Specificity | 40.7                | 31.5               |
|                         | NPV         | 100                 | 100                |
|                         | PPV         | 59.0                | 55.4               |
|                         | Az value    | 0.8660              | 0.8770             |
| 3rd year 1              | Sensitivity | 97.8                | 97.8               |
|                         | Specificity | 48.2                | 51.9               |
|                         | NPV         | 96.3                | 96.6               |
|                         | PPV         | 61.6                | 63.4               |
|                         | Az value    | 0.8770              | 0.8730             |
| 3rd year 2              | Sensitivity | 87.0                | 95.7               |
|                         | Specificity | 59.3                | 57.4               |
|                         | NPV         | 84.2                | 93.9               |
|                         | PPV         | 64.5                | 65.7               |
|                         | Az value    | 0.8080              | 0.8830             |
| Total resident reviewer | Sensitivity | 94.2                | 97.8               |
|                         | Specificity | 50.7                | 40.0               |
|                         | NPV         | 92.1                | 95.9               |
|                         | PPV         | 62.1                | 56.4               |
|                         | Az value    | 0.8543              | 0.8695             |
| Expert                  | Sensitivity | 97.8                |                    |
|                         | Specificity | 64.8                |                    |
|                         | NPV         | 97.2                |                    |
|                         | PPV         | 70.3                |                    |
|                         | Az value    | 0.9440              |                    |

NPV: negative predictive value, PPV: positive predictive value, Az value: area under curve



**Fig. 3.** Longitudinal (A) and transverse (B) image. Two reviewers thought this lesion was iso echogenic so they categorized this lesion as category 3 before training, but they changed their minds to complex echogenic, and changed the category of this lesion to 4a. Experienced breast radiologists decided this lesion had complex internal echogenicity and categorized it as category 4a. This lesion was malignant. In this case, accuracy decreased after training, but they learned the complex echogenicity as the feature which increases the possibility of malignancy. Agreement was improved after training.

(31.5-51.9%)로 감소를 보였다. 6명중 1명에서만 특이도가 증가하는 경향을 보였으나 통계적인 차이는 없었다. 6명 중 4명에서 음성 예측도가 증가하였고, 3명에서 양성 예측도가 증가하였다. Az value는 6명의 평가자중 4명에서 (평균, 0.0248; 범위, 0.0020- 0.0750) 증가하였으나 통계적 차이는 없었다 ( $p > 0.05$ ).

### 고 찰

현재까지 유방촬영술이 유방암의 선별을 위한 가장 유용한 검사방법이며 민감도는 약 85%까지 보고되고 있으나, 한국 여성과 같이 유방밀도가 치밀한 경우에 있어서는 민감도가 감소하게 된다 [9]. 유방초음파 검사는 최근 고주파수의 탐촉자가 개발되고 해상도가 향상됨에 따라서 유방촬영술의 보조적인 진단방법으로서의 역할 이외에도 치밀유방에 있어서 유방암의 선별검사로서 매우 중요한 역할을 하게 되었다. 그러나 이러한 영상검사가 유방암의 선별검사 및 진단 검사로서 유용한 역할을 하기 위해서는 양질의 영상을 얻을 수 있어야 하며, 적절한 해석을 할 수 있어야 한다. 이러한 이유로 ACR에서 유방촬영술 및 유방초음파검사에 대한 BI-RADS를 발행하였으며 [10-12], 이는 유방보고를 표준화하고, 유방 영상해석의 혼동을 줄이기 위해 고안된 정도 관리방법이다. 또한 미국식품의약국 (Food and Drug administration, FDA)은 유방촬영술 정도 관리 표준화법 (Mammography Quality Standards Act)를 제정하고 유방촬영술을 관독하는 의사의 자격을 규정하였으며, 이는 반드시 영상의학과 전문의여야 하며, 처음으로 독립적인 관독의사가 되기 전에 직접감독 하에 약 240여건의

유방촬영술을 관독한 경험이 있어야 하며, 전공의가 유방촬영술을 관독 시엔 전문의와 버금가는 조건이 필요함을 규정하고 있다 [2, 13]. 그러나 유방초음파는 유방촬영술에 비해서 전공의가 실제로 해 볼 수 있는 기회가 적으며 많은 경우가 전문의들에 의해 이루어 지고 있고 [1], 유방초음파를 관독하는데 있어서 필요한 구체적인 교육의 양과 자격에 관한 규정은 아직 제정된 바는 없다.

유방초음파 검사는 실시간 검사이므로 촬영한 의사가 제대로 촬영하지 못하면 병변을 해석할 수 없다. 적절한 장비와 병변의 정확한 영상을 촬영하는 것이 기본이 되어야 하며 유방촬영술과 연관을 지어 해석하는 것이 매우 중요하다. 초음파 BI-RADS도 유방촬영술 BI-RADS의 목적과 같이 통일된 용어를 사용함으로써 영상의학과 의사, 임상사간의 의견 교환을 원활히 하며 나아가서 국가적인 통계자료의 기본을 마련하는데 그 의의가 있다 [11]. 그러나 초음파 검사의 가장 큰 단점 중 하나는 영상 관독의 관독자간 (inter-observer), 관독자 내의 변이 (intra-observer variation)가 많아 객관적이 판단이 어렵다는 것이다. 또한 구체적인 영상소견과 최종 평가 범주와 연관시킨 규정이 없기 때문에 최종판단 및 용어 사용에 있어 초음파를 시행하는 의사의 경험이 매우 중요하다. 그러므로 경험이 부족한 검사자가 검사를 할 경우 경험 많은 사람이 검사했을 때에 비해 병변에 대한 해석이 달라지거나 병변을 발견하는 율 자체도 상이 할 수 있다는 단점이 있다. 이러한 단점을 극복하기 위해 영상의학과 전공의가 유방초음파 검사에 처음 노출되어 관독하는데 있어 적절한 교육이 필요하며, 유방초음파 BI-RADS 어휘사용과 최종 평가 범주를 결정하는데 있어 교육의 역할이 크다고 할 수 있다.

본 연구에서, 병변의 해석에 있어서 종괴 내부 에코와 병변 가장자리 항목에서 가장 많은 교육효과가 있었다. 이는 교육을 통하여 내부 에코 평가의 기준이 주변의 지방 (fat) 이라는 점을 알게 된 결과로 생각하며, 병변 가장자리 항목에서는 테두리로 보이는 부분과 주변 유방 실질의 고에코성 부분의 구별 방법을 배웠다는 것을 보여주었다. 교육 후에 방향성과 변연 항목에서는 통계적으로 유의하지는 않지만, 일치도가 감소하는 경향을 보였다. 하지만 방향성 항목은 교육 전부터 높은 일치도를 보여서 이미 연구 이전에 이루어진 다른 분야의 초음파 경험과 강의를 통해서 교육이 어느 정도 이루어졌다는 것을 의미한다. 또 변연 항목은 교육 전 후에 모두 일치도가 0.4이하로 낮은 값을 보였는데, 이는 1달 또는 2달의 유방 초음파 경험과 교육으로는 전문의 수준의 해석을 하기 어렵다는 결과로 볼 수 있고, 유방 초음파에서 변연 항목 평가에 있어서는 가장 많은 경험과 교육이 필요하다는 것을 의미한다. 전공의들은 1달 간의 교육을 통해, 전반적인 민감도가 증가하였으며, 전문평가자들과의 비교에 있어서도 일치도가 다소 증가한 양상을 보였다. 특히 2년차 전공의의 경우 평가 항목 중 41.7% (15/36) 에서 통계적으로 유의한 일치도의 증가를 보여 18개 항목 중 단 3개의 항목에서 일치도의 유의한 증가를 보인 고년차 전공의에 비해, 높은 일치도 향상률을 보였다. 특히 2년차 평가자 1에서는 가장 많은 항목에서 변화를 보여서 교육의 효과가 가장 컸다. 이는 교육 전 일치도의 값이 다른 전공의들보다 낮은 경향이 있어서 그 차이가 더 많이 나타난 것으로 생각된다. 교육 후 전공의들의 일치도 값은 교육 전보다 개별적인 차이가 줄어든 것을 볼 수 있어 교육에 의해서 전공의들의 수준이 어느 정도까지 향상되었다는 것을 알 수 있었다. 그러나, 1달간의 수련기간을 통해 민감도 면에선 전공의와 유방영상 전문의 간의 격차를 줄일 수 있었지만, 특이도 면에선 오히려 떨어지는 양상을 보였으며, 전체적인 Az value에선 큰 차이를 보이지 않았다. 특이도를 포함한 전반적인 ROC 값의 증가를 얻기 위해서는 오랜 기간의 경험이 필요하리라 예상되며, 이에 대해서는 추가적인 연구가 필요할 것으로 생각된다.

본 연구의 제한점은 연구의 대상들이 유방영상 및 유방초음파에 있어 연구 기간 이전에 교육받은 바가 있다는 것이다. 연구의 대상 중 2년차 들은 이전에 유방 초음파에 대한 경험이 없지만 유방학에 대한 지식을 강의 등을 통하여 얻은 적이 있고, 3년차의 경우에는 이전에 1달간의 유방 초음파 경험이 있다. 또한 모든 대상자가 복부, 갑상선 등, 다른 장기의 초음파를 시행 판독한 경험들이 수개월씩 있어 다른 부분이긴 하지만 초음파를 시행하면서 초음파의 특성을 이해하고, 악성과 양성 병변의 감별점을 익혔기 때문에, 이들 경험이 유방 초음파의 영상을 해석 하는데 영향을 줄 수 있다. 교육 받은 전공의 들 3명이 교육 전 평가에서 이미 민감도가 95% 이상을 보여, 이 연구에 참여한 전공의의 상당수가 이미 교육된 상태임을 보여주는 결과라 할 수 있다. 그래서 연구기간 전에 연구자들이 이미 소유한 지식의 양으로 인하여 교육 전 후의 평가 결과의 차

이가 줄어들었을 수 있었을 것으로 생각한다. 그러므로, 유방영상학과 초음파에 전문적인 지식 및 교육이 없이 새로이 유방초음파를 시작하는 의사들에 있어서 교육의 기간과 교육 정도에 대한 연구의 결과는 차이가 있을 수 있고, 이에 대한 추가적인 연구가 필요하다. 또한 만약 신입 전공의를 대상으로 했다면 유방영상 및 초음파를 처음 시작하는 의사에 대한 연구가 가능하리라 생각하며, 교육 전 후의 차이가 더 컸을 것이다. 따라서, 전반적인 전공의의 필요한 교육기간과 정도를 구하기 위해서는 신입전공의를 포함한 다양한 경험의 차이를 보이는 다수의 전공의를 포함한 연구가 필요하다.

본 연구는 정지된 사진으로 평가하였으며, 임상적인 소견과 초음파에서 보이는 병변 간에 긴밀한 영향이 있으나, 교육내용에는 이들의 연관성이 포함되지 않았다. 마지막으로 본 연구는 1달의 수련기간 전 후의 평가만을 시행하였으며, 수련 후 수개월 후에 수련 내용이 얼마나 유지되는지에 대해서는 포함하지 않아서, 이에 대한 추가 연구가 필요할 것으로 생각한다.

본 연구에선 유방 초음파의 교육을 위해, 전공의에게 전문의 감독 하에 200 여건 이상의 실시간 유방 초음파 검사 및 판독, 100 여건 이상의 유방 촬영술 판독, 유방초음파검사에 대한 BI-RADS를 탐독과 약 50개의 유방초음파 사례에 대해 전문평가자가 판독한 BI-RADS 어휘와 최종평가범주를 본인의 것과 비교, 그리고 전문의에 의해 진행되는 유방 영상학에 관한 강의를 2시간 이상의 교육을 실시하였다. 앞으로 전공의기간 동안 유방영상학의 교육을 위한 수련 기간, 수련을 위해 적절한 검사 건수, 및 교육의 방법에 대해선 추가적인 연구가 필요할 것으로 생각한다.

결론적으로, 본 연구는 유방초음파에 대한 전공의 기간 동안의 교육은 유방 초음파에서 보이는 병변을 해석해서 최종 평가를 내리는데 결과를 향상시키며, 전문 평가자가 평가한 내용과 일치도 향상에 도움이 됨을 보였으며, 교육의 방법, 기간 등에 대해서는 추가적인 연구가 필요할 것으로 생각한다.

## References

1. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR Am J Roentgenol* 2003;181:177-182
2. Feig SA, Hall FM, Ikeda DM, et al. Society of Breast Imaging residency and fellowship training curriculum. *Radiol Clin North Am* 2000;38:915-920, xi
3. Bassett LW, Monsees BS, Smith RA, et al. Survey of radiology residents: breast imaging training and attitudes. *Radiology* 2003;227:862-869
4. Berg WA, D'Orsi CJ, Jackson VP, et al. Does training in the Breast Imaging Reporting and Data System (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography? *Radiology* 2002;224:871-880

5. Nodine CF, Kundel HL, Mello-Thoms C, et al. How experience and training influence mammography expertise. *Acad Radiol* 1999;6:575-585
6. Fleiss JL. *Statistical methods for rates and proportions*. Wiley New York, 1981
7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174
8. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 1971;76:378-382
9. Rosenberg RD, Hunt WC, Williamson MR, et al. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. *Radiology* 1998;209:511-518
10. D'Orsi CJ, Basset LW, Feig SA. *Breast Imaging Reporting and Data System (BI-RADS)*. Reston, VA: American College of Radiology 1998
11. Mendelson EB, Baum JK, Berg WA, Merritt CRB, Rubin E. *Illustrated Breast Imaging Reporting and Data System (BI-RADS): ultrasound*. Reston, VA: American College of Radiology 2003
12. American College of Radiology. *Breast Imaging Reporting and Data System (BI-RADS)*. VA: Reston, 2003;229-251
13. Sickles EA, Philpotts LE, Parkinson BT, et al. American College of Radiology/Society of Breast Imaging Curriculum for Resident and Fellow Education in Breast Imaging. *Journal of the American College of Radiology* 2006;3:879-884



= Abstract =

## The Influence of Training on Breast Ultrasound Expertise

**Song-Ee Baek, M.D., Min Jung Kim, M.D., Eun-Kyung Kim, M.D.,  
Eun-Hee Choi, M.S.\*, Hye Ryoung Koo, M.D., Ji Youn Kim, M.D.,  
Honsoul Kim, M.D., Sung Ho Park, M.D., Jung Hyun Yoon, M.D.,  
Mi-Jung Lee, M.D., Seung Woo Lee, M.D., Jin Young Kwak, M.D.**

Department of Radiology, Research Institute of Radiological Science, Yonsei University College of Medicine

\*Department of Biostatistics, Yonsei University College of Medicine

**PURPOSE :** This study was designed to evaluate the effectiveness of the training experience to develop expertise for breast ultrasound. The same group of radiology residents was evaluated before and after training on the use of the Breast Imaging Reporting and Data System (BI-RADS) to develop breast imaging expertise.

**MATERIALS and METHODS :** An evaluation set of breast ultrasonography images of pathologically confirmed lesions that consisted of 100 masses (54 benign and 46 malignant masses) without clinical information was developed. Two experienced breast radiologists independently reviewed cases to reach a consensus. Six radiology residents (four second-year residents and two third-year residents) trained one month to interpret breast ultrasound examinations, and to read and to review teaching files with expertise. The residents evaluated the examination set according to BI-RADS for shape, orientation, margin, lesion-boundary, echo pattern, posterior echo features, surrounding tissue, calcification and final assessment before and after the training. Training performance was measured by means of agreement (k) and diagnostic performance.

**RESULTS :** After training, the second-year residents showed statistically significant increased agreement as compared to the experts for 15 of the total 36 BI-RADS descriptors. The third-year residents showed increased agreement for three of the 18 descriptors. The mean sensitivity for the residents was 94.2% (range, 89.1–100%) before training and the mean sensitivity increased to 97.8% (range, 95.7–100%) after training. The mean specificity was 50.6% (range, 40.7–59.3%) before training and the mean specificity decreased to 42.0% (range, 31.5–51.9%) after training.

**CONCLUSION :** In this study, the use of breast ultrasonography training improved the ability of residents to determine lesion description, final assessment and coherence.

Address for reprints : Min Jung Kim, M.D., Department of Radiology, Research Institute of Radiological Science, Yonsei University College of Medicine, 250 Seongsanno, Seodaemun-gu, Seoul 120-752, Korea.  
Tel. 82-2-2228-7400 Fax. 82-2-393-3035 E-mail: mines@yuhs.ac