

유방촬영술 인공지능은 검진 목적의 보조 초음파를 대체할 수 있을까?

장정민

서울대학교병원 영상의학과

유방촬영술은 유방암 조기 발견과 사망률 감소에 기여해 온 표준 검진 방법이지만, 치밀 유방에서는 민감도가 감소하는 한계가 있다. 최근 딥러닝 기반 인공지능(Artificial Intelligence, AI)의 발전으로 유방촬영술 판독 정확도가 향상되었고, 일부 연구에서는 단독 또는 이중 판독을 대체할 가능성도 제시되고 있다. 그러나 유방촬영술에서 보이지 않고 초음파로만 발견되는 작은 비석회성 암은 유방촬영술 AI 기반 분석에서 검출이 어려울 수 있어, AI가 검진 목적의 보조 초음파를 완전히 대체할 수 있는지에 대해서는 여전히 의문이다. 본 종설에서는 유방촬영술 AI와 보조 초음파의 진단 성적을 비교한 연구들을 고찰하고, AI의 한계와 초음파의 대체 및 병합 사용의 가능성을 논의하고자 한다.

Index words: Mammography, Dense breasts, Artificial intelligence, Supplemental ultrasound

서론

유방촬영술은 유방암 조기 발견을 위한 검진 도구로 자리 잡아왔으며, 대규모 집단 검진 프로그램을 통해 유방암 사망률 감소에 기여하는 것으로 알려져 왔다 (1). 그러나 치밀 유방에서는 민감도가 감소하는 한계가 있어 이를 보완하고자 유방초음파, 삼차원 디지털 토모신테시스 또는 MRI 등이 활용되고 있다 (2,3). 특히 아시아 여성에서 치밀 유방의 빈도가 높다는 점은 초음파를 활용한 검진의 임상적 역할을 더욱 부각시켜 왔다.

최근 인공지능(Artificial Intelligence, AI)의 비약적인 발전은 유방촬영술 판독 환경에 중대한 변화를 가져오고 있다. 딥러닝 기반 알고리즘은 병변 검출, 악성도 예측, 판독

보조 측면에서 우수한 성능을 보고하고 있으며, 일부 연구에서는 단독 판독 또는 이중 판독을 대체할 가능성도 제시하고 있다 (4). 이러한 기술적 진보는 치밀 유방에서의 민감도 한계를 극복할 수 있을 것이라는 기대를 낳고 있다. 그러나 유방촬영술 AI가 보조 초음파를 완전히 대체할 수 있는지에 대해서는 여전히 논쟁의 여지가 있다. 초음파로 발견된 암들은 그 크기가 작고 석회를 동반하지 않는 경우가 많아 후향적으로 다시 유방촬영술을 확인하더라도 찾기 어려운 경우가 많고, 치밀 유방의 경우 더욱 확인이 어렵다. 유방암이 보이는 불규칙한 모양, 침상변연 또는 동반된 구조왜곡이 결여된 경우 유방촬영술 AI로 병변을 진단하는 것에도 한계가 있을 수 있다.

본 종설에서는 유방촬영술 AI 활용과 부가적인 초음파의 성적을 직간접적으로 비교한 연구들을 바탕으로 유방촬영술 AI의 현재 성과와 한계를 고찰하고, 검진 목적의 보조 초음파 검사의 대체 및 병용사용의 필요성에 대하여 알아보고자 한다.

통신저자: 장정민
(03080) 서울시 종로구 대학로 101
서울대학교병원 영상의학과
Tel. 82-2-2072-2584 Fax. 82-2-747-7418
E-mail: imchangjm@gmail.com

치밀유방에서의 유방촬영술 인공지능의 활용

최근 유방촬영술 분야에서의 딥러닝 및 인공지능(AI) 발전은 유방암 진단에 있어 영상의학과 전문의 수준과 비등한 판독을 가능하게 하였으며, 치밀 유방 조직에 숨어있는 암의 발견율을 향상시키는 동시에 위양성 결과를 감소시키고 판독 효율성을 개선할 잠재력을 보이고 있다. 디지털 유방촬영 검진에서 단독 AI의 성능을 평가한 최근 메타분석에 따르면, AI는 개별 유방 영상의학과 전문의 또는 평균 수준의 판독자들의 성적과 동등하거나 그 이상의 성능을 보였다. 또한 AI를 활용한 유방촬영 판독은 치밀 유방과 비치밀 유방 여성 모두에서 진단 성능의 향상을 보이는 것으로 보고되었다 (4). 유방암 검진을 위한 유방촬영술을 시행한 국내 여성 89,855명을 대상으로 AI단독 판독과 영상의학과 전문의의 유방촬영술 판독 성능을 비교한 대규모 후향적 연구에 따르면, 유방 밀도에 따른 성적을 비교하였을 때, 영상의학과 전문의와 AI 모두 비치밀 유방(non-dense)에서 가장 좋은 성능을 보였다. Category C의 heterogeneously dense breast에서는 전문의의 암 발견율과 민감도가 AI 보다 다소 높았으나 통계적으로 뚜렷하지는 않았던 반면, 특이도 93.0% [95% CI, 92.9-93.2] vs. 77.6% [95% CI, 61.7-77.9]), 양성예측도 (1.5% [95% CI, 1.2-1.9] vs. 0.5% [95% CI, 0.4-0.6]), 재검률 (7.1% [95% CI, 6.9-7.2] vs. 22.5% [95% CI, 22.2-22.7])은 모든 밀도 범주에서 AI가 일관되게 우수했고 (모든 $P < 0.05$). 특히 category D, extremely dense breast에서 가장 두드러지게 AI가 좋은 성적을 보였다 (5).

유방촬영술 인공지능이 놓치는 암의 특성

AI가 치밀유방을 포함한 모든 유방 밀도에서 우수한 성적을 보이기는 하나, AI가 유방촬영술에서 모든 암을 찾을 수 있는 것은 아니다. 한 연구에 따르면 (6), 1097개의 유방암을 AI를 통해 분석한 결과 14% (154건)에서 AI가 암을 진단하지 못했다. 유방암 분자아형별로 위음성률을 비교 분석하였을 때에는 HER2-enriched 아형에서 9%로 가장 낮았고, Luminal 아형은 17.2%, 삼중음성 아형은 14.5%로 통계적으로 유의한 차이를 보였다 ($P = 0.001$). AI가 검출한 암과 비교했을 때, AI가 놓친 암은 더 젊은 환자에서 발생했고, 종양 크기가 2cm 이하로 작으며, 조직학적 등급이 낮고, 림프절 전이가 적고, Ki-67 발현이 낮은 특성을 보였다(모두 $P < 0.05$). 또한 BI-RADS 4 범주에 해당하는 경우가 더 많았고, 유방의 변연부에 위치

한 경우가 많았다. 초음파로 발견된 것으로 기록된 253개의 유방암의 유방촬영술 영상을 후향적으로 AI를 적용하여 분석한 또 다른 연구에서, AI는 66개의 암 (26.1%)을 검출한 반면, 187개의 암 (73.9%)은 여전히 검출하지 못하였다 (7). 특이한 점은 AI가 검출한 66개의 암 중 42개 (63.6%)는 치밀 유방 조직에 가려져 명확히 구분하기 어려운 병변이었다는 것이다. 이러한 결과는 AI가 치밀 유방에서 사람이 쉽게 인지하지 못하는 미묘한 특징을 일부 포착할 수 있음을 시사한다. 국내 다기관 전향연구로 진행중인 AI for Breast Cancer Screening in Mammography (AI-STREAM) 연구 자료를 이용하여 AI가 단독으로 찾거나, AI만 놓치는 암을 분석한 결과에 따르면, 치밀 유방에서 미미한 소견을 보이거나 Luminal A 아형에 해당하는 유방암은 사람이 판독에서 놓친 경우에도 AI로 검출되는 사례가 있었다. 그러나 반대로, 사람 판독자가 발견했지만 AI가 놓친 경우도 치밀 유방의 경우가 많았다 (8). 결국 치밀 유방에서 명확한 이상 소견이 없는 작은 암의 경우에는 판독자나 AI 모두에서 검출에 어렵다는 것을 알 수 있다 (9).

치밀 유방에서의 유방촬영술 인공지능 활용과 보조 초음파의 성적 비교

치밀 유방여성에서 AI와 추가적인 보조 초음파의 성적을 직접적으로 비교한 연구로 다음의 두 연구가 있다 (표 1). 한 연구에서는 1325명의 치밀 유방을 가진 여성을 대상으로 유방촬영술 단독, AI, 초음파 단독성적과 유방촬영술+AI, 유방촬영술+초음파 성적을 비교하였다 (10). 총 1325명에서 12건의 유방암이 진단되었다. 단독 검사 비교에서 유방촬영술, AI, 초음파 간 암 발견율과 민감도는 유의한 차이가 없었으나, AI와 초음파는 유방촬영술보다 재검률이 높고 특이도와 정확도가 낮았다. 병합 검사 비교에서는 유방촬영술과 초음파를 함께 시행한 경우가 유방촬영술과 AI를 병합하거나 유방촬영술, AI, 초음파 세 가지를 모두 병합한 경우보다 특이도와 정확도가 높고 재검률이 낮았다. 또한 유방촬영술과 초음파 병합 시 민감도는 100%로 가장 높았다. 전반적으로 치밀 유방 환자에서 보조 초음파를 시행한 상황에서는 AI를 추가하더라도 뚜렷한 추가 이득은 확인되지 않았다 (10). 또 다른 연구에서도 검진기관 여성들 중 치밀 유방을 가진 무증상 여성 5707명을 대상으로 유방촬영술 단독, 유방촬영술+AI, 그리고 유방촬영술+보조 초음파의 진단 성능을 비교하였다 (11). 총 33건 (0.6%)의 유방암이 진단되었으며, 병변의 중앙 크기는 0.7cm였다. 유방촬영술+AI는 유방촬영

표 1. 유방촬영술과 AI병합 및 유방촬영술과 초음파 병합시 성적 비교

	Lee et al. AJR 2024 (1,325 women)			Ha et al. Radiology 2024 (5,707 women)		
	Mammography with AI	Mammography with US	P value	Mammography with AI	Mammography with US	P value
CDR*	7.5 (3.6–13.8) [10/1,325]	9.1 (4.7–15.8) [12/1,325]	0.67	3.5 (2.3, 5.4) [20/5,707]	5.6 (3.9, 7.9) [32/5,707]	0.002
Recall Rate	14.9 (13.0–16.9) [197/1,325]	11.7 (10.0–13.6) [155/1,325]	0.02	5.0 (4.5, 5.6) [286/5707]	22.9 (21.8, 24.0) [1,305/5,707]	<0.001
Sensitivity	83.3 (51.6–97.9) [10/12]	100.0 (73.5–100.0) [12/12]	0.12	60.6 (43.7, 75.3) [20/33]	97.0 (84.7, 99.5) [32/33]	0.002
Specificity	85.8 (83.7–87.6) [1,126/1,313]	89.1 (87.3–90.7) [1,170/1,313]	0.009	95.3 (94.7, 95.8) [5,408/5,674]	77.6 (76.5, 78.6) [4,401/5,674]	<0.001

Note. Except where indicated, data are percentages, with 95% CIs in parentheses and numbers of examinations in brackets.

* Data are reported as the rate per 1000 examinations. CDR = cancer detection rate

술 단독보다 특이도가 유의하게 높고 (95.3% vs 94.3%), 낮은 비정상 판독률(Abnormal interpretation rate, AIR)을 보였다 (5.0% vs. 6.0%). 그러나 암 발견율과 민감도 측면에서는 뚜렷한 향상을 보이지 않았다. 반면, 유방촬영술+초음파는 유방촬영술+AI에 비해 암 발견율(5.6 vs 3.5 per 1000)과 민감도(97.0% vs 60.6%)가 유의하게 높았으며, 초음파 검사는 유방촬영술+ AI 가 찾지 못한 12건의 암을 발견했고 이 중 92%가 0기 또는 1기의 림프절 음성 조기암이었다. 다만, 초음파 병합 시 특이도는 낮아지고 (77.6%) AIR 높아졌다(22.9%). 결론적으로 AI는 부가적으로 활용하는 것은 유방촬영술의 특이도를 향상시켰지만, 조기 림프절 음성 유방암을 더 많이 발견하는 데에는 초음파를 병합하는 것이 더 효과적이었다 (11). 두 연구에서 보이는 각종 진단 성적 지표의 차이는 연구에 포함된 여성의 집단의 차이, 즉 검사 적응증, 유방 밀도 등으로 생각되나, 두 연구 모두에서 AI 가 찾지 못하는 암을 초음파는 찾아낼 수 있음을 공통적으로 확인할 수 있었다.

인공지능을 활용한 보조 초음파 검사 대상자의 발굴

위와 같은 연구 결과들을 고려할 때 유방촬영술 AI는 초음파를 대신할 수 없으며, 영상의학과 의사의 유방촬영술 판독이 정상이고, AI 도 이상 소견 없음으로 보고하는 경우에도 유방암이 있을 수 있음을 알 수 있다. 그러나, 정상 유방촬영술, AI 소견에서 추가적인 암이 나오는 빈도는 매우 낮다는 것과 유방촬영술을 이용한 유방암 위험도 예측이 가능할 수 있다는 연구들이 등장함에 따라 모든 여성이 초음파 검사를 하는 것이 아닌 보조 초음파 검사를 해서 득을 볼 수 있는 여성, 즉 추가적인 암을 찾을 수 있는 여성을 유방촬영술 AI를 통해 알아내는 방법이 있을지 고민해볼 수

있다.

그 예로 Liu 등(12)은 암 소견 탐지, 가림 효과 (masking effect) 평가, 유방암 발생 위험을 추정하는 세 가지 AI 모델을 간소하게 결합한 모델인 AISmartDensity를 개발하여 정상 유방촬영술 이후 보조 유방 영상검사의 대상자를 선별하는 데 있어 AISmartDensity 점수의 유용성을 평가하였다. AISmartDensity는 2008년 1월부터 2015년 12월까지 스웨덴 인구 기반 코호트에서 시행된 65,325건의 유방촬영술을 대상으로 검증되었으며, 해당 코호트는 암 특성 정보가 포함된 암 등록 자료와 연계되어 있었다. 정상 유방촬영술 결과를 받은 환자에서 향후 유방암 발생을 예측하는 성능을 기존의 전통적인 유방 밀도 모델과 비교한 결과 통합 AI 모델은 기존 밀도 기반 모델보다 우수한 성능을 보였으며, AISmartDensity의 AUC는 0.72로, 가장 우수한 밀도 모델의 AUC 0.61보다 유의하게 높았다 ($P < 0.001$) (12). 이러한 연구는 아직 초기 단계로 제한된 환자군과 일부의 소프트웨어를 이용한 연구만 진행되었고, 임상에 어떻게 적용될 수 있는지에 대한 방안에 대해서도 많은 추가 연구와 검증이 필요한 상황이다. 향후 정상 유방촬영술 결과를 얻었으나 보조 검진 검사를 시행하면 좋겠다고 생각되는 여성을 AI를 이용하여 정확하게 가려낼 수 있다면 해당 여성에게서만 초음파나 MRI 검사를 보다 효율적으로 시행하는 검진이 시행될 수도 있을 것이다.

결론

유방촬영술 AI 기술은 앞으로도 더 발전할 것이고 인간의 눈으로 인지하지 못하는 유방암을 찾아내는 민감도는 더욱더 높아질 것이다. 그러나, 유방촬영술이 가지는 검사 자체의 한계 및 초음파 또는 MRI 등 다른 영상 기법이 가

지는 장점으로 기인한 추가적인 암의 발견은 계속 존재할 수밖에 없다. 현재까지의 근거를 종합하면, 유방촬영술 기반 AI는 판독의 정확도 향상과 재검률 감소 측면에서 유방촬영술 단독보다는 의미 있는 장점을 보이지만, 치밀 유방에서 보조 초음파를 완전히 대체하기는 어렵다. 유방촬영술 AI를 활용하여 유방촬영술 진단능을 향상시키고, 보조 초음파를 적절히 보완적으로 활용한다면 유방암 검진의 효율을 높일 수 있을 것이라 기대한다. 또한, 향후 대규모 전향적 연구와 실제 임상 환경에서의 비용-효과 분석이 축적된다면, AI 기반 선별 전략에 기반한 보조 초음파의 활용 등이 가능할 것으로 생각된다.

References

1. Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *Lancet*. 2003; 361(9367): 1405 – 1410.
2. Hooley RJ, Greenberg KL, Stackhouse RM, Geisel JL, Butler RS, Philpotts LE. Screening US in patients with mammographically dense breasts: initial experience with Connecticut Public Act 09-41. *Radiology*. 2012; 265(1):59-69.
3. Vourtsis A, Berg WA. Breast density implications and supplemental screening. *Eur Radiol*. 2019; 29(4):1762-1777.
4. Yoon JH, Strand F, Baltzer PAT, et al. Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: a systematic review and meta-analysis. *Radiology*. 2023; 307(5):e222639.
5. Kwon MR, Chang Y, Ham SY, et al. Screening mammography performance according to breast density: a comparison between radiologists versus standalone intelligence detection. *Breast Cancer Res*. 2024; 26(1):68.
6. Woo OH, Song SE, Choe SJ, Kim M, Cho KR, Seo BK. Invasive breast cancers missed by AI screening of mammograms. *Radiology*. 2025; 315(3):e242408.
7. Yoen H, Chang JM. Artificial intelligence improves detection of supplemental screening ultrasound-detected breast cancers in mammography. *J Breast Cancer*. 2023; 26(5):504-513.
8. Chang YW, Ryu JK, An JK, Choi N, Park YM, Ko KH. Breast cancers detected and missed by AI-CAD: results from the AI-STREAM trial. *Radiol Artif Intell*. 2026; 8: e250281
9. Chang JM. What AI-CAD detects – and misses – in real-world breast cancer screening. *J Korean Soc Radiol*. 2026; 87(1):52-53.
10. Lee SE, Yoon JH, Son NH, Han K, Moon HJ. Screening in patients with dense breasts: comparison of mammography, artificial intelligence, and supplementary ultrasound. *AJR Am J Roentgenol*. 2024; 222(1):e2329655.
11. Ha SM, Jang MJ, Youn I, et al. Screening outcomes of mammography with AI in dense breasts: a comparative study with supplemental screening US. *Radiology*. 2024; 312(1):e233391.
12. Liu Y, Sorkhei M, Dembrower K, Azizpour H, Strand F, Smith K. Use of an AI score combining cancer signs, masking, and risk to select patients for supplemental breast cancer screening. *Radiology*. 2024; 311(1):e232535.