

유방외과 의사와 S-Detect™의 유방종괴에 대한 진단수행도 비교

전주예수병원 외과

김하람, 권의영, 박영삼, 최은혜, 김미진, 김철승

A Comparative Study of the Diagnostic Performance of Evaluating Breast Masses for Breast Surgeons versus S-Detect™ (Samsung Medison Co., Ltd, Seoul, Korea)

Haram Kim, Eiyoungh Kwon, Youngsam Park, Eunhye Choi, Mijin Kim, Cheolseung Kim

Department of Surgery, Presbyterian Medical Center, Jeonju, Korea

Received August 9, 2019
Revised October 7, 2019
Accepted November 1, 2019**Purpose:** Ultrasonography is widely used for examining breast mass. We used the Breast Imaging-Reporting and Data System (BI-RADS) to characterize breast lesions found on ultrasonography. Among various ultrasound techniques, we used S-Detect™ (Samsung Medison Co., Ltd, Seoul, Korea), which supports the morphological analysis of breast masses found according to BI-RADS. In addition, we compared the breast surgeons' categorization of breast masses with that by S-Detect™.**Methods:** Breast surgeons evaluated the breast masses found using ultrasonography between April 2016 and December 2016. A total of 139 masses, which were categorized as BI-RADS 3 or 4, from 112 patients were reevaluated by S-Detect™ before performing vacuum-assisted resection or surgical excision.**Results:** Of the 139 masses, 118 were benign tumors and 21 were malignant tumors. With regard to the diagnostic performance, the sensitivity of categorization was 95% for breast surgeons, but the sensitivity was relatively lower for S-detect™ (85%). However, the specificity and accuracy of S-detect™ were 70.6% and 74.1%, respectively, which were higher than those values obtained from breast surgeons (18.5% and 30.9%, respectively).**Conclusion:** S-detect™ can be used by breast surgeons as a diagnostic aid when evaluating and diagnosing breast masses found on ultrasonography.**Keywords:** Breast, Ultrasonography, Computer-assisted diagnosis, Artificial intelligence**Correspondence to:**
Cheolseung Kim
Department of Surgery,
Presbyterian Medical Center, 365
Seowon-ro, Wansan-gu, Jeonju
54987, Korea
Tel: +82-63-230-1408
Fax: +82-63-230-1409
E-mail: cskimmd@hotmail.com

서론

유방암은 여성 암 사망의 가장 흔한 원인이며, 또한 이로 인한 사망률은 점차 증가하고 있다. (1) 유방암의 5년 생존률은 유방암의 병기가 낮을수록 사망률이 더 낮은 경향

을 보이는데 2018년 한국유방암학회의 발표에 따르면 조기 유방암으로 분류되는 0, 1, 2기 환자에서는 90% 이상의 5년 생존율을 보이거나 전신 전이가 있는 4기 환자는 34%로 상대적으로 낮은 생존율을 보였다. (1) 따라서 유방암의 생존률을 높이기 위해서는 무엇보다도 유방암을 조기에

발견하는 것이 중요하다. 유방암의 선별검사로 유방촬영술이 쓰이고 있지만 우리나라에서는 50세 미만의 여성에서 치밀 유방의 비율이 높아 유방암으로 진단된 환자의 90% 이상이 유방초음파를 통해 진단되고 있어 유방초음파의 역할이 크다.(2) 하지만 유방초음파로 관찰된 유방 병변을 유방암으로 조기에 확진하고 악성여부가 불명확한 병변의 경우에 조속히 악성여부를 감별하기 위해서는 조직검사를 적극적으로 시행할 수밖에 없다.(3)

조직검사의 필요성을 명확하게 하기 위한 초음파 판독 기준으로 BI-RADS (Breast imaging reporting and data system)가 사용되고 있다. BI-RADS는 초음파시술자와 유방외과 전문의 간에 유방 병변에 대한 표준화된 기준을 제공하여 의사결정에 도움을 준다.(4) 유방초음파상 악성의 가능성이 있는 소견은 불분명한 경계, 미세엽상형, 불규칙 모양, 불균질 내부 에코, 혈류 증가 소견, 주변부 고음영, 각진 변연, 석회화, 유선관 확장 등이 있다.(4) 하지만 표준검사 소견이 있다 하더라도 양성과 악성 종괴의 초음파 소견들은 부분적으로 중복될 수 있으며,(4) 또 시술자에 따라 유방 병변의 판독에 대한 민감도가 떨어질 수 있다. 이러한 단점을 보완하기 위해 인공지능기반 진단 보조 프로그램이 등장하였고, S-detect™ (Samsung Medison Co.,Ltd, Seoul, Korea), CADstream (Merge Healthcare Inc., Chicago, IL, USA), 및 DynaCAD for Breast (Invivo, Gainesville, FL, USA) 등이 사용되고 있다.(5) 그 중에서 S-detect™는 초음파로 측정되는 종양의 형태를 인식하여 평가하는 방식을 취하고 있다.

이 연구는 유방외과 전문의가 평가한 BI-RADS와 S-detect™의 종괴 분석평가 결과를 비교하고, 절제 후 확인된 병리조직검사와 각각을 비교 분석하여 유방암에 대한 진단 기기로서의 S-detect™의 성능을 평가하고자 하였다.

방 법

이 연구는 후향적 연구로, 본원의 IRB 승인을 받았다(No. 2019-07-026).

1. 환자군

2016년 4월부터 12월까지 유방 종괴를 주소로 본원 외래를 방문한 환자들 중에 유방초음파상 BI-RADS cat-

egory 3 또는 4에 해당하는 종괴를 갖는 환자들은 121명이었다. 본 연구는 미세석회화(microcalcification) 3명, 사진이 저장되지 않은 환자 3명, 과거 섬유선종(fibroadenoma) 제거 후 재발한 1명, 중심부바늘생검(core-needle biopsy)을 시행하지 않고 추적관찰 하기 원한 환자 2명 등 9명을 제외한 112명을 대상으로 하였다. 총 112명의 환자에서, 합병 낭종(complicated cystic mass) 양상으로 중심부바늘생검을 시행하지 않고 바로 진공 보조 흡인 생검술을 시행한 종괴 5개와 중심부바늘생검을 시행한 134개의 종괴를 합하여 총 139개의 종괴를 평가하였다.

2. 종괴의 평가

유방 외과 전문의 3명이 유방 종괴를 1차적으로 평가하였으며, BI-RADS category에 따라 분류하였다. BI-RADS category 3; probably benign; 4a, low suspicion for malignancy; 4b, intermediate suspicion for malignancy; 4c, moderate concern for malignancy에 해당하는 유방 종괴에 대하여 초음파 유도 하 중심부 바늘 생검을 시행하여 양성으로 확인된 종괴는 진공 보조 흡인 생검술을 시행하였고, 악성 종양으로 확인된 경우 수술적 치료를 통해 제거하였다. 진공 보조 흡인 생검술과 수술을 시행하기 전에 S-detect™를 이용하여 종괴를 재평가하였다. S-detect™는 종괴의 깊이(depth), 너비(width), 높이(height), 모양(shape), 방향성(orientation), 변연(margin), 병변의 경계(lesion boundary), 후방 특성(posterior feature), 에코 양상(endo pattern), 주변 조직(surrounding tissue), 혈관 분포 상태(vascularity), 특별 증례(special case) 등을 분석하여 possibly benign 또는 possibly malignant로 유방 종괴를 분류하였다.

3. 통계

Medical calc version 18.0 (Panmun education, Seoul, Korea)을 이용하였다. 양성 종괴군과 악성 종괴군의 크기 비교 및 유방 종괴 크기에 따른 초음파와 병리조직검사의 일치여부 상관관계를 분석하고 각 환자군의 평균 나이와 1 cm 이상 크기를 가지는 종괴 비율에 대한 통계적 분석을 위하여 T-test를 시행하였다. BI-RADS와 S-detect™의 결과를 비교하기 위해서 Generalized estimating equation method를 적용하였다. 또한 BI-RADS 및 S-detect™의 Receiver Operating Characteristic

(ROC) curve 비교를 위해 Delong's method를 사용하여 평가하였다.

결 과

전체 139개의 종괴 중에서 양성 종양은 118개(84.2%)였으며, 악성 종양은 21개(15.8%)였다. 연구대상 환자들의 평균나이는 44.4 ± 0.25 세였으며, 유방 종괴의 평균 크기는 12.1 ± 0.47 mm였다. 유방 종괴 중에서 양성 종괴의 평균 크기는 14.7 ± 7.55 mm였으며, 악성 종괴의 평균 크기는 11.6 ± 0.47 mm였다. 양성 유방병변 중 1 cm 미만은 52개, 1 cm 이상은 66개였고, 악성 유방병변 중 1 cm 미만은 7개, 1 cm 이상은 14개였다(Table 1).

유방종괴를 병리조직검사를 바탕으로 분류한 결과는 Table 2와 같다. 양성 종양 중에서는 섬유선종(Fibroadenoma)이 가장 많았고, 악성 종양 중에서는 침윤성 유관암(invasive ductal carcinoma)가 가장 많았다.

Fig. 1과 2는 S-detect™를 이용하여 유방 종괴를 평가한 결과이다. Fig. 1은 타원형 모양의 경계가 분명한 종괴를 S-detect™가 Possibly benign으로 평가한 것을 보여

준다. 이 종괴의 조직병리검사 결과는 섬유선종으로 나왔다. Fig. 2에서는 불규칙한 모양의 경계가 불분명한 종괴를 S-detect™가 Possibly malignant로 평가하였으며, 최종 병리검사 결과 침윤성 유관암으로 확인되었다.

유방외과 전문의에 의해 측정된 BI-RADS category에 대해서 S-detect™를 통해 재평가한 결과를 반영하여 분류하였다(Table 3). 병리조직검사에서 악성을 보인 유방 종괴 중에서 IVc로 분류된 유방 종괴가 S-detect™의 Possibly malignant로 측정된 비율이 가장 많았지만 통계적으로 유의하지는 않았다($P = 0.954$).

유방외과 전문의와 S-detect™의 초음파 결과에 대한 진단 수행도(diagnostic performance)를 비교하였다(Table 4). BI-RADS III를 benign으로, IVa, IVb, IVc를 malignant로 각각 설정하여 비교하였을 때, 유방전문의 민감도(sensitivity)는 95%, NPV 95.652%로 S-detect™보다 더 높은 결과를 보였고 통계적으로 유의하였다($P <$

Table 1. Patient Characteristics

	Benign n = 118	Malignancy n = 21	P-value
Age (years)	38.8 ± 2.31	45.5 ± 1.01	0.009
Tumor size (mm)	14.7 ± 7.55	11.6 ± 0.47	0.065
< 1 cm (number)	52 (44.1%)	7 (33.3%)	0.175
≥ 1 cm (number)	66 (55.9%)	14 (66.7%)	0.175

Expressed as mean \pm standard deviation.

Table 2. Pathologic Diagnosis

Diagnosis	Pathology	No. (%)
Benign (n = 118, 84.9%)	Fibroadenoma	76 (64.1)
	Fibrocystic change	13 (11.1)
	Adenosis	12 (10.3)
	Benign fibroepithelial lesion	8 (6.8)
	Fluid ductal hyperplasia	4 (3.4)
	Intraductal papilloma	2 (1.7)
	Sclerosing adenosis	2 (1.7)
	Sclerosing lobular hyperplasia	1 (0.8)
	Invasive ductal carcinoma	19 (90.5)
Malignancy (n = 21, 15.1%)	Ductal carcinoma in situ	1 (4.8)
	Invasive lobular carcinoma	1 (4.8)

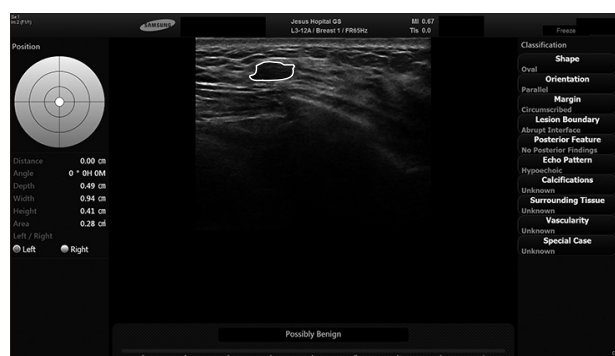


Fig. 1. A 27-year-old woman with left breast mass on B-mode ultrasound. S-Detect™ concluded that it is "Possibly Benign" based on features of the lesion listed on the right column.

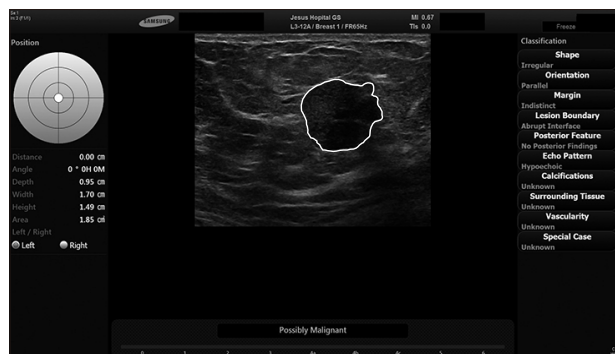


Fig. 2. A 65-year-old woman with left breast mass on B-mode ultrasound. S-Detect™ concluded that it is "Possibly Malignant" based on features of the lesion listed on the right column.

Table 3. Distribution of Final Assessments of the 139 Breast Masses according to the Breast Surgeons and S-Detect™

BI-RADS by surgeons	Pathologic results	S-detect™			P-value
		Possibly benign	Possibly malignant	Total	
III	Benign	17	5	23	0.219
	Malignant	1	0		
IVa	Benign	54	23	84	0.073
	Malignant	1	6		
IVb	Benign	13	5	19	0.379
	Malignant	0	1		
IVc	Benign	0	1	13	0.954
	Malignant	1	11		

BI-RADS = Breast Imaging-Reporting and Data System.

Table 4. Diagnostic Performance for Breast Surgeons versus S-Detect™

	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	AUC
BI-RADS	95	18.5	30.9	16.4	95.7	0.567
S-detect™	85	70.6	74.1	32.69	96.6	0.778
P-value	0.0055	< 0.0001	< 0.0001	0.0016	0.6988	< 0.0001

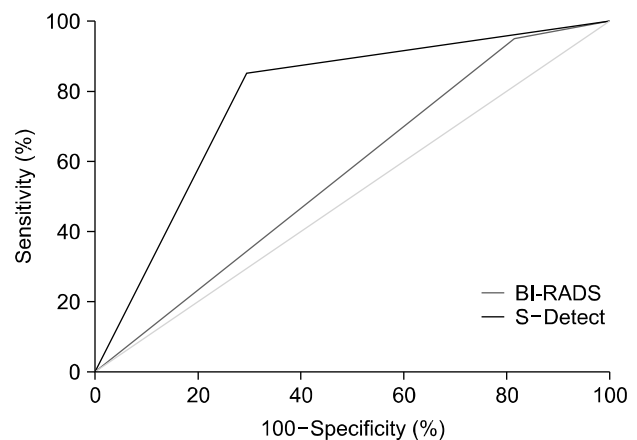
BI-RADS = Breast Imaging-Reporting and Data System; PPV = positive predictive value; NPV = negative predictive value; AUC = area under the receiver operating characteristic curve.

Table 5. Mean Size of the Masses according to the Results of BI-RADS and S-Detect™ and Histopathological Results

	Benign (n = 119)	Malignant (n = 20)	P-value
BI-RADS3	15.09 mm (22)	7 mm (1)	0.008
BI-RADS4	11.13 mm (97)	13.85 mm (19)	
S-detect™ possibly benign	11.19 mm (84)	9.86 mm (3)	0.168
S-detect™ possibly malignant	13.46 mm (35)	14.15 mm (17)	

0.05). S-detect™ for Breast의 경우에는 특이도 (specificity), 정확도(accuracy), 양성 예측도(PPV), AUC에서 더 높은 결과를 보였으며, 모두 통계적으로 유의하였다. 초음파 결과에 대한 Receiver Operating Characteristic (ROC) curve는 Fig. 3과 같다.

유방 종괴에 대하여 유방외과 전문의와 S-detect™가 평가한 초음파 결과와 조직병리검사 결과가 유방 종괴 크기가 클수록 일치하는 경향이 있는지 상관관계를 분석하였다. Table 5 같이, S-detect™와 달리 유방외과 전문의는 유방 종괴가 클수록 병리조직검사 결과와 통계적으로 유의하게 더욱 일치하는 경향이 있음을 보였다(P = 0.008).

**Fig. 3.** Receiver operating characteristic curve for breast surgeons (BI-RADS) and S-Detect™.

고 찰

현재 우리나라의 유방검진은 만40세 이상에서 2년마다 정기적으로 유방촬영술을 통해 시행할 것을 권고하고 있다. 하지만 유방의 치밀 정도가 유방촬영술의 민감도를 감소시키는 가장 중요한 요인으로 유방 치밀도가 높은 우리나라에서는 유방촬영술에서 발견하지 못한 조기 유방암을 발견하기 위한 검사수단으로서 유방초음파의 역할이 크다. (6-8) 초음파는 인체에 무해한 장점이 있지만 초음파를 시행하는 의사의 술기 능력에 따라 진단의 정확성이 달라진다. 초음파 술기가 미흡한 경우 병변에 대한 위양성 판정률이 상승되어 불필요한 조직검사 등의 추가적인 검사 및 치료가 이어질 수 있다. (9)

위양성률을 줄일수 있는 첫번째 방법은 BI-RADS를 이용하여 평가하는 것이다. 하지만 유방초음파를 습득하는 과정이 쉽지 않고 BI-RADS의 정확한 기술이 시술자의 숙련도에 크게 영향을 받는 점과 외관상 양성병변과 악성병변이 겹쳐지는 경우 기술이 쉽지 않은 단점이 있다. (10) 두 번째로는 인공지능을 통해 병변을 다방면으로 측정하여 보다 정확하게 악성여부를 판단하고 이상소견을 놓치지 않는 것이다. (11,12)

S-detect는 앞서 언급한 것처럼 종괴의 깊이, 너비, 높이, 모양, 방향성, 변연, 병변의 경계, 후방 특성, 에코 양상, 주변 조직, 혈관 분포 상태, 특별 증례 등을 분석하여 possibly benign 또는 possibly malignant로 평가하여 초음파 시술자에게 악성여부를 제시해준다. 하지만 최종적인 판단은 초음파를 시행하는 의사가 결정한다. (13)

실제로, Choi 등(14)은 경험이 있는 영상의학과 전문의가 초음파를 통해 BI-RADS를 이용하여 유방 종괴를 평가하는 것보다 S-detectTM가 정확도, 양성 예측도에서 통계적으로 유의하게 성적이 좋았다고 하였으며, 경험이 적은 의사의 경우 S-detectTM를 이용할 때 오진(misdiagnosis)의 빈도가 줄어든다고 하였다. 그리고 Kim 등(15)은 S-detect가 영상의학과 전문의보다 특이도, 양성 예측도, 정확도, AUC에서 통계적으로 유의하게 좋은 결과를 보였다고 하였고, S-detect를 이용하면 특이도를 향상시킬 수 있다고 하였다. 또한 Shin 등(16)은 유방 종괴의 크기가 1.2 cm 이상인 경우 유방 종괴의 자동화된 평가 기술을 신뢰할 수 있다고 하였다.

이번 연구에서도 삼성 메디슨의 S-detectTM를 이용하

여 인공지능의 진단 도움이 유용한지에 대해서 알아보았다. 본 연구 결과에서 민감도는 유방외과 전문의를 통한 BI-RADS 평가가 95%로 85%의 결과를 보인 S-detectTM보다 더 높은 결과를 보였지만 특이도(specificity), 정확도(accuracy), 양성 예측도(PPV), AUC의 경우에는 유방외과 전문의보다 S-detectTM에서 통계적으로 유의하게 높은 결과를 보여주었다. 그리고 유방외과 전문의 경우에는 크기가 클수록 병변의 병리조직검사와 더 일치하는 경향을 보였다.

이러한 S-detectTM의 특성으로 인해 S-detectTM를 이용하여 유방 종괴를 평가했던 이전의 연구들을 보면, S-detectTM를 이용하여 종괴를 평가한 경우가 S-detectTM없이 초음파를 시행했을 때보다 평가 성적이 더 좋아지는 결과를 살펴볼 수 있다. (15,16) 그렇지만 이번 연구에서도 볼 수 있듯이 S-detectTM의 결과를 절대적으로 신뢰하기에는 아직 부족한 면이 있다. 초음파 특성상 시술자의 능력에 영향을 받을 수 있기에 유방초음파를 시행하는 의사의 충분한 기본 소양의 함양이 무엇보다도 가장 필수적이겠으며, S-detectTM가 추가적으로 진단에 보조적인 역할로 사용된다면 더욱 좋은 결과를 보여줄 것으로 여겨진다.

결 론

최근 인공지능 기술이 급속히 발전하면서 의학에도 영향을 미치고 있으며 특히 질병의 진단을 돕는 데 많이 쓰이고 있다. S-detectTM 또한 이러한 장치 중 하나로서, 우리가 초음파를 이용하여 유방 병변을 진단할 때, 보조장치로 충분히 활용할 수 있을 것으로 생각된다.

REFERENCES

1. Korean Breast Cancer Society. Breast Cancer Facts & Figures 2018. Seoul: Korean Breast Cancer Society; 2018.
2. Youn I, Choi S, Kook SH, Choi YJ. Mammographic breast density evaluation in Korean women using fully automated volumetric assessment. J Korean Med Sci 2016;31:457-62.
3. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995;196:123-34.

4. Kim YJ, Choi HY, Moon BI, Lee SN. Categorization and evaluation of usefulness of breast lesions with using ultrasound BI-RADS (Breast Imaging Reporting and Data system). *J Korean Radiol Soc* 2006;54: 313–8.
5. Pan J, Dogan BE, Carkaci S, Santiago L, Arribas E, Cantor SB, et al. Comparing performance of the CADstream and the DynaCAD breast MRI CAD systems : CADstream vs. DynaCAD in breast MRI. *J Digit Imaging* 2013;26:971–6.
6. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002;225:165–75.
7. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology* 2001;221:641–9.
8. Brem RF, Lenihan MJ, Lieberman J, Torrente J. Screening breast ultrasound: past, present, and future. *AJR Am J Roentgenol* 2015;204:234–40.
9. Kopans DB. Sonography should not be used for breast cancer screening until its efficacy has been proven scientifically. *AJR Am J Roentgenol* 2004;182:489–91.
10. Galperin M, Andre MP, Barker CH, Olson LK, O'Boyle M, Richman K, et al. Reproducibility of image analysis for breast ultrasound computer-aided diagnosis. *Acoustical Imaging* 2009;29:397–402.
11. Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis, and Machine Vision*. 4th ed. Boston: Cengage Learning; 2014.
12. Chen CM, Chou YH, Han KC, Hung GS, Tiu CM, Chiou HJ, et al. Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks. *Radiology* 2003;226:504–14.
13. Fujita H, Uchiyama Y, Nakagawa T, Fukuoka D, Hatanaka Y, Hara T, et al. Computer-aided diagnosis: the emerging of three CAD systems induced by Japanese health care needs. *Comput Methods Programs Biomed* 2008;92:238–48.
14. Choi JH, Kang BJ, Baek JE, Lee HS, Kim SH. Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience. *Ultrasonography* 2018;37:217–25.
15. Kim K, Song MK, Kim EK, Yoon JH. Clinical application of S-Detect to breast masses on ultrasonography: a study evaluating the diagnostic performance and agreement with a dedicated breast radiologist. *Ultrasonography* 2017;36:3–9.
16. Shin HJ, Kim HH, Cha JH, Park JH, Lee KE, Kim JH. Automated ultrasound of the breast for diagnosis: interobserver agreement on lesion detection and characterization. *AJR Am J Roentgenol* 2011;197: 747–54.