Medical auditing of whole-breast screening ultrasonography

Min Jung Kim

Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

Since breast ultrasonography (US) has been used as an adjunctive screening modality in women with dense breasts, the need has arisen to evaluate and monitor its possible harm and benefits in comparison with other screening modalities such as mammography. Recently, the fifth edition of the Breast Imaging Reporting and Data System published by the American College of Radiology has suggested auditing methods for screening breast US. However, the method proposed therein is slightly different from how diagnostic performance was calculated in previous studies on screening breast US. In this article, the background and core aspects of medical audits of breast cancer screening will be reviewed to provide an introduction to the medical auditing of screening breast US, with the goal of helping radiologists to understand and identify potential ways to improve outcomes.

Keywords: Breast neoplasms; Ultrasonography; Screening; Medical audit

Introduction

Mammography is the only screening modality proven to reduce breast cancer deaths among women aged 40 years and older in previous randomized clinical trials [1], despite questions regarding the degree of its contribution in improving breast cancer survival in screening settings [2]. However, the false-negative rates of mammography can reach 50% in dense breasts [3], and the survival benefit of performing mammography is lower in younger women with dense breasts [4], suggesting the necessity of developing adjunctive screening modalities. One of the suggested adjunctive modalities is ultrasonography (US), for which the incremental cancer detection rates have been reported to be 2.7 to 4.2 per 1,000 women screened [5,6]. However, several aspects of breast US need to be verified before it can be used as the primary screening test for breast cancer.

Although improved survival is the main goal of breast cancer screening programs, obtaining follow-up data and evaluating the final outcomes requires a lengthy period of time. Therefore, researchers have identified several surrogate markers that can be assessed within a short-term follow-up and that are also periodically used to monitor the efficacy of screening programs. Medical audits for breast cancer screening programs with mammography have been well established [7,8], along with legislation making such audits mandatory [9]. However, the medical auditing of screening breast US has not been established to date.

In this article, the background and core aspects of medical audits of breast cancer screening

are reviewed to provide an introduction to the medical auditing of screening breast US, with the goal of helping radiologists to understand and identify potential ways to improve outcomes.

Background of the Medical Auditing of Screening Mammograms

The main purpose of a medical audit, first described in a white paper of the U.K. National Health Service in 1989, is to improve patient care and outcomes, and the main elements of medical audits are the systematic review of care against solid criteria, the implementation of changes, and monitoring to confirm improvement [10,11]. For mammography, the purpose of a medical audit is to provide feedback to facilities and doctors on their performance relative to established benchmarks and to improve the overall quality of the breast cancer screening program. Several reports have shown the effectiveness of medical auditing. One showed improved sensitivity, with maintained positive predictive values and an increased number of detected cancers, of which most were node-negative, during 2 years of education and feedback on the audit results were provided to radiologists [12]. These results showed improvements in the effectiveness of the screening program after the medical audit, despite a 50% increase in the use of additional mammographic views and sonography. The Breast Imaging Reporting and Data System (BI-RADS) was established by the American College of Radiology (ACR) in response to communication issues and societal concerns about false negatives occurring in the context of a nationwide screening mammography program [13–15]. The ACR organized the overall structure of mammography reports, including assessment categories and the corresponding management strategies, supplied benchmarks for mammographic interpretation, and furthermore, developed the National Mammography Database within the National Radiology Data Registry [15].

Medical audits are also necessary because of one serious harm of mammography that has recently become an issue: anxiety among healthy women due to positive results, of which most will turn out to be false positives, with benign results on biopsy [16,17]. It has been found that two or three breast cancers will be identified among 100 women who receive annual mammograms over the course of 10 years. However, of these 100 women, 62 will eventually have positive results on their mammograms. Furthermore, seven to 10 will undergo biopsies with benign pathology results. Therefore, it must be determined whether the anxiety arising from the 62 false-positive results and 10 benign biopsies is inevitable and/or whether it is justifiable given the two or three detected breast cancers. Searching for ways to lower the excess amount of false-positive results and benign biopsies needs to be a part of the medical audit.

The management of screening mammography should be a constant balance between possible harm and benefits; in other words, between missing too many cancers and ordering too many biopsies and recalls. This balance should be attained through medical auditing [9,15,18].

Currently, the following basic elements of a medical audit are recommended: (1) to track all positive mammograms, (2) to correlate the imaging findings with the pathologic results of all biopsies performed, and (3) to review known false-negatives within 12 months of the mammography examination [19]. The majority of approximately 250 radiologists in a survey on medical audits agreed that audit reports were valuable [20]. However, approximately half of the radiologists with less than 10 years of experience and three-fourths of radiologists in their early 30s replied that they would consider ceasing to interpret mammograms if congress mandated more intensive auditing requirements without providing funds to support the regulation. Thus, conducting more extensive medical audits needs both the cooperation of radiologists and the provision of sufficient resources [10,20].

Characteristics and Audit Data of Screening Breast US: Comparison with Benchmarks for Mammography

Cancer Detection Rate

The role of screening breast US in dense breasts has been studied since the early 2000s, and was recently further emphasized by new legislation in the United States requiring patients to be notified of their breast density after a screening mammography [21]. In asymptomatic women, supplementary hand-held US performed by radiologists detected 1.8–5.3 additional cancers per 1,000 women [5,21–24], and the overall cancer detection rate in supplemental screening US was 3.7 per 1,000 women (248 of 66,828) in 10 single- or multi-center studies [25]. Among these asymptomatic women, based on known risk factors, more than three-fifths of women with breast cancers detected only by screening US (89 of 145 women) were found to be at elevated risk in seven studies [5,6,18,21,26–29]. Recently, since the implementation of the Connecticut Public Act, breast US has been performed on average-risk women as well, and the cancer detection rate of breast US for these women (1.6–2.4 per 1,000 examinations) [21,24,30,31] has tended to be lower than the range accepted for screening mammography (≥2.5 per 1,000 examinations) [32] and the previously observed detection rate in high-risk women (3.7–5.3 per 1,000 examinations) [5]. As more women without risk factors for breast cancer undergo screening breast US, the acceptable range of the cancer detection rate needs to be reconsidered for screening

breast US [31].

Characteristics of Breast Cancers Detected by Screening US

Supplemental breast US improved the detection of mostly node-negative invasive breast cancer in past studies; approximately 90% of breast cancers detected by screening US were stage 0 or I and node-negative [25], a much higher percentage than the benchmarks for mammography (74.8% and 77.3%, respectively) [32], which suggests that breast US can detect early breast cancer. However, the positive predictive value of the recommendation for tissue diagnosis (PPV$_2$) and the positive biopsy rate (PPV$_3$) showed a large proportion of benign pathologic results (3.3%–11.7%), much lower than the acceptable range for mammography (20%–40%) [5,21,22,24,31]. This suggests that further studies are needed on US interpretation and supplemental sonographic techniques to improve the positive predictive value of screening breast US.

The majority of breast cancers detected by breast US (over 90%) are known to be invasive, which may suggest that breast US can detect early but life-threatening breast cancer [5,33]. However, further studies must be performed to generalize these results, because the data were also based on study populations with a large proportion (70%–100%) of high-risk women, as was the case for the cancer detection rates. Recent studies focusing on women with an average risk for breast cancer showed much lower rates of the detection of invasive cancer (range, 0% to 60%; overall mean, 50% [5 of 10]) (Table 1) [21,22,24,34]. Moreover, these results

are also somewhat different from the results of studies on high-risk women or screening mammography. In particular, the audit data for screening breast US among women at an average risk for breast cancer demonstrated a higher abnormal interpretation rate (25.0%–31.1% vs. 20.9% and 11.5%, respectively), a lower positive predictive value for biopsies (3.4%–5.3% vs. 9.0% and 29.2%, respectively), and more common recommendations for short-term follow-up (22.3%–26.6% vs. 11.1%–13.5% and 3.2%, respectively). These disparities indicate that it is necessary to collect various benchmarks for screening breast US according to risk stratification.

Proportion of probably benign findings

An abundant proportion of BI-RADS 3 (probably benign) findings were found to require additional short-interval follow-up in studies on screening US [5,21,22,24]. Because BI-RADS recommends not using BI-RADS 3 (probably benign) findings in the interpretation of screening mammography, the data for breast US do not correspond to what is expected. Although breast US is used as a screening modality, it includes a diagnostic component as well as a screening component. In fact, no guidelines exist regarding the proportion of BI-RADS 3 findings in diagnostic mammography. Therefore, it is inevitable that a higher proportion of BI-RADS 3 lesions are found through breast US screening than through screening mammography. However, recommending additional examinations prior to the next screening examination results in patient anxiety and demands

Table 1. Summary of studies on screening breast ultrasonography according to the risk for breast cancer in comparison with results for screening mammography

Variable	Berg et al. [5], 2012		Hooley et al. [21], 2012		Chang et al. [22], 2015	Moon et al. [24], 2015	Weigert and Steenbergen [34], 2015
Modality	MG	US	US		US	US	US
Risk	Increased	Increased	Increased	Average	Average	Average	Not specified
No. of examinations	7,473	7,473	236	614	1,526	2,005	10,282
No. of cancers	59	58	2	1	5	4	24
Yield, per 1000	7.9	7.7	8.5	1.6	3.3	2.0	2.3
Invasive	41/59 (69.5)	53/58 (91.4)	2/2 (100)	0/1 (0)	3/5 (60.0)	2/4 (50.0)	10/21 (47.6)[a]
Node-negative	18/41 (43.9)	34/53 (64.2)	2/2 (100)	–	3/3 (100)	2/2 (100)	–
Abnormal inter-pretation (recall) rate	306/2,659 (11.5)[b]	555/2,659 (20.9)[b]	234 (25.0)		431 (28.2)	623 (31.1)	1,310 (12.7)
Short-term follow-up	84/2,659 (3.2)[b]	296/2,659 (11.1)[b]	About 32 (13.5)	About 155 (25.3)[c]	340 (22.3)	533 (26.6)	875 (9.0)
Biopsy rate	65/2,659 (2.4)[b]	233/2,659 (8.8)[b]	66/935 (7.1)		91/1,526 (6.0)	88/2,005 (4.4)	435/10,282 (4.2)
PPV$_3$	19/65 (29.2)[b]	21/233 (9.0)[b]	3/66 (4.5)		5/91 (5.3)	3/88 (3.4)	24/435 (5.5)[d]

Values are presented as number (%).

MG, mammography; US, ultrasonography; PPV$_3$, positive biopsy rate.

[a]Two carcinomas (invasive/uncertain) and one lymphoma were excluded [34]. [b]Calculated based on a prevalence screen [5]. [c]Inferred from Fig. 2 of the Hooley et al.'s article [21]. [d]Eight point nine percent if high-risk pathologic lesions are included (39 of 435) [34].

additional resources. This could be a hurdle in the implementation of screening breast US. Therefore, we need to collect data and establish benchmarks for the proportion of BI-RADS 3 (probably benign) findings in screening breast US.

Most of the audit data on whole-breast screening US reviewed so far have been obtained from previous studies that focused on asymptomatic women who had negative mammograms in dense breasts, in whom screening US was used as a supplemental modality after mammographic information was already obtained, with the exception of the American College of Radiology Imaging Network (ACRIN) 6666 study [5,23]. Therefore, we need to recognize that the currently available data cannot be utilized to evaluate the use of US alone to screen for breast cancer.

Medical Audit of Breast US Suggested by ACR BI-RADS

The definition of a positive screening result in breast US has been continuously debated [35,36]. Based on the principle that its definition should be objective and reproducible, and that comparisons with other modalities should be possible, a positive screening was defined by the ACR as the acquisition of additional recorded images such as additional mammographic views, regardless of the duration of the examination or whether a second opinion was requested [32]. Breast US also includes a diagnostic component as well as a screening component. A single report may be issued for both components because it would be awkward to report the two components of a single US examination separately. However, the ACR suggested separate auditing for the screening and diagnostic components of combined US examinations [32].

According to the protocol of a previous multicenter study, which was proposed as a benchmark study in BI-RADS, the image requirement for a negative examination is at least five images per

breast, one for each quadrant, and one for the retroareolar area [37]. If a nonbenign breast lesion is found in the screening setting, the examination is considered to have become diagnostic. Therefore, the screening breast US report should indicate whether only standard images were recorded or whether additional (diagnostic) images were obtained. Exams with additional images should be considered audit-positive for the screening component, regardless of the final assessment of breast US (Table 2). Therefore, a positive screening includes all subcategories of BI-RADS categories 3, 4, and 5, and categories 1 and 2 with more than five images per breast. Negative screenings include only BI-RADS categories 1 and 2 with five images per breast. Additionally, orthogonal images are not allowed for BI-RADS category 2, whereas all BI-RADS categories 1 and 2 findings were considered as negative screenings regardless of the number of images in the previous literature [5,6,18,21,22,24,26,28,30,31,33,34]. Therefore, the medical audit data will change in correspondence with the ACR recommendations.

Potential Ways to Improve the Outcomes of Medical Audits of Breast US

As discussed above, a high short-term follow-up recommendation rate and low positive predictive value are characteristics of screening breast US [38]; however, several studies have reported reducing the number of BI-RADS categories 3 and 4 cases by applying supplemental techniques, such as shear-wave elastography (SWE) and/or Doppler US [39,40], or by applying strict criteria to the imaging findings [41,42]. The addition of SWE and Doppler US to B-mode US increased the specificity by more than 45% points in non-mass lesions detected by screening US without loss in sensitivity (from 10 of 42 to 29 of 42) and helped 65% of patients with BI-RADS category 4a lesions (19 of 29) avoid unnecessary biopsies [39]. By applying the strict criteria of the ACRIN 6666 protocol to imaging

Table 2. Reference for interpreting biopsy results (screening and diagnostic) according to the ACR BI-RADS

Interpretation		According to the BI-RADS, fifth edition					
		5 Images for each breast		≥6 Images for either breast			
		Biopsy-positive	Biopsy-negative	Biopsy-positive		Biopsy-negative	
Component: screening (S) or diagnostic (D)		S	S	S	D	S	D
BI-RADS categories 1, 2	Image results	Negative		Positive	Negative	Positive	Negative
	Audit data	FN	TN	TP	FN	FP	TN
BI-RADS category 3	Image results	–	–	Positive	Negative	Positive	Negative
	Audit data	–	–	TP	FN	FP	TN
BI-RADS categories 4, 5	Image results	–	–	Positive	Positive	Positive	Positive
	Audit data	–	–	TP	TP	FP	FP

ACR, American College of Radiology; BI-RADS, Breast Imaging Reporting and Data System; FN, false negative; TN, true negative; TP, true positive; FP, false positive.

findings, 18.3% of BI-RADS category 3 lesions (213 of 1,164) were retrospectively recategorized as category 2 lesions [42]. Over the course of 3 years after education and feedback on audit results were provided in clinical practice, the rate of BI-RADS categories 3 to 4a was halved (from 22.6%–39.4% to 11.1%–16.0%) and the biopsy rates also significantly decreased, from 6.5% to 2.4%, while the cancer detection rate of supplemental screening US was maintained (2.8 per 1,000 examinations) [41].

Conclusion

Medical audits are essential for quality assurance. Therefore, they are not optional for breast cancer screening programs, and surrogate measures for medical audits of screening breast US need to be adjusted and developed. Medical audits will help screening breast US to become widely and effectively used as the primary screening test with the goal of improving the survival of breast cancer patients.

ORCID: Min Jung Kim: http://orcid.org/0000-0003-4949-1237

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghate S, et al. Benefits and harms of breast cancer screening: a systematic review. JAMA 2015;314:1615-1634.
2. Harding C, Pompei F, Burmistrov D, Welch HG, Abebe R, Wilson R. Breast cancer screening, incidence, and mortality across US counties. JAMA Intern Med 2015;175:1483-1489.
3. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081-1087.
4. Nelson HD, Cantor A, Humphrey L, Fu R, Pappas M, Daeges M, et al. Screening for breast cancer: a systematic review to update the 2009 U.S. Preventive Services Task Force Recommendation. Rockville, MD: Agency for Healthcare Research and Quality, 2016.
5. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA 2012;307:1394-1404.
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;225:165-175.
7. Farria DM, Monsees B. Screening mammography practice essentials.

Radiol Clin North Am 2004;42:831-843.
8. Sickles EA. Quality assurance: how to audit your own mammography practice. Radiol Clin North Am 1992;30:265-275.
9. United States General Accounting Office. Mammography Quality Standards Act: X-ray quality improved, access unaffected, but impact on health outcomes unknown. GAO/T-HEHS-98-164. Washington, DC: US General Accounting Office, 1998.
10. Shaw CD, Costain DW. Guidelines for medical audit: seven principles. BMJ 1989;299:498-499.
11. Shaw CD. Medical audit in Britain: what now and what next? Qual Assur Health Care 1989;1:61-63.
12. Linver MN, Paster SB, Rosenberg RD, Key CR, Stidley CA, King WV. Improvement in mammography interpretation skills in a community radiology practice after dedicated teaching courses: 2-year medical audit of 38,633 cases. Radiology 1992;184:39-43.
13. Moss M. Spotting breast cancer: doctors are weak link. New York Times 2002 Jun 27:Sect. A1, A22.
14. Scott WC. Establishing mammographic criteria for recommending surgical biopsy. Report of the Council on Scientific Affairs. Chicago, IL: American Medical Association, 1989.
15. Burnside ES, Sickles EA, Bassett LW, Rubin DL, Lee CH, Ikeda DM, et al. The ACR BI-RADS experience: learning from history. J Am Coll Radiol 2009;6:851-860.
16. Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. Ann Intern Med 2011;155:481-492.
17. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of breast cancer screening: systematic review to update the 2009 U.S. Preventive Services Task Force Recommendation. Ann Intern Med 2016;164:256-267.
18. Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dunser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. Semin Ultrasound CT MR 2000;21:325-336.
19. U.S. Food and Drug Administration. Quality standards and certification.s90012 [Internet]. Silver Spring, MD: U.S. Food and Drug Administration, 2017 [cited 2017 Jan 17]. Available from: http://www.fda.gov/radiation-emittingproducts/mammographyqual itystandardsactandprogram/regulations/ucm110906.htm.
20. Elmore JG, Aiello Bowles EJ, Geller B, Oster NV, Carney PA, Miglioretti DL, et al. Radiologists' attitudes and use of mammography audit reports. Acad Radiol 2010;17:752-760.
21. 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;265:59-69.
22. Chang JM, Koo HR, Moon WK. Radiologist-performed hand-held ultrasound screening at average risk of breast cancer: results from

a single health screening center. Acta Radiol 2015;56:652-658.

23. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008;299:2151-2163.

24. Moon HJ, Jung I, Park SJ, Kim MJ, Youk JH, Kim EK. Comparison of cancer yields and diagnostic performance of screening mammography vs. supplemental screening ultrasound in 4394 women with average risk for breast cancer. Ultraschall Med 2015;36:255-263.

25. Merry GM, Mendelson EB. Update on screening breast ultrasonography. Radiol Clin North Am 2014;52:527-537.

26. 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.

27. Gordon PB, Goldenberg SL. Malignant breast masses detected only by ultrasound: a retrospective review. Cancer 1995;76:626-630.

28. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. Radiology 2001;221:641-649.

29. Berg WA. Supplemental screening sonography in dense breasts. Radiol Clin North Am 2004;42:845-851.

30. Chae EY, Kim HH, Cha JH, Shin HJ, Kim H. Evaluation of screening whole-breast sonography as a supplemental tool in conjunction with mammography in women with dense breasts. J Ultrasound Med 2013;32:1573-1578.

31. Parris T, Wakefield D, Frimmer H. Real world performance of screening breast ultrasound following enactment of Connecticut Bill 458. Breast J 2013;19:64-70.

32. Sickles EA, D'Orsi CJ. ACR BI-RADS follow-up and outcome monitoring. In: D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA, eds. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System.

Reston, VA: American College of Radiology, 2013;1-67.

33. Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol 2010;20:734-742.

34. Weigert J, Steenbergen S. The connecticut experiments second year: ultrasound in the screening of women with dense breasts. Breast J 2015;21:175-180.

35. Berg WA, Mendelson EB. How should screening breast US be audited? The patient perspective. Radiology 2014;272:309-315.

36. Sickles EA, D'Orsi CJ. How should screening breast US be audited? The BI-RADS perspective. Radiology 2014;272:316-320.

37. Berg WA, Mendelson EB. Technologist-performed handheld screening breast US imaging: how is it performed and what are the outcomes to date? Radiology 2014;272:12-27.

38. Moon HJ, Kim EK. Characteristics of breast cancer detected by supplementary screening ultrasonography. Ultrasonography 2015;34:153-156.

39. Choi JS, Han BK, Ko EY, Ko ES, Shin JH, Kim GR. Additional diagnostic value of shear-wave elastography and color Doppler US for evaluation of breast non-mass lesions detected at B-mode US. Eur Radiol 2016;26:3542-3549.

40. Lee SH, Chang JM, Kim WH, Bae MS, Seo M, Koo HR, et al. Added value of shear-wave elastography for evaluation of breast masses detected with screening US imaging. Radiology 2014;273:61-69.

41. Kim SY, Kim MJ, Moon HJ, Yoon JH, Kim EK. Application of the downgrade criteria to supplemental screening ultrasound for women with negative mammography but dense breasts. Medicine (Baltimore) 2016;95:e5279.

42. Chae EY, Cha JH, Shin HJ, Choi WJ, Kim HH. Reassessment and follow-up results of BI-RADS category 3 lesions detected on screening breast ultrasound. AJR Am J Roentgenol 2016;206:666-672.