Diagn Interv Radiol 2023; DOI: 10.5152/dir.2022.211047

BREAST IMAGING

ORIGINAL ARTICLE

Prediction of malignancy upgrade rate in high-risk breast lesions using an artificial intelligence model: a retrospective study

Özge Aslan
Ayşenur Oktay
Başak Katuk
Rıza Cenk Erdur
Oğuz Dikenelli
Levent Yeniay
Osman Zekioğlu
Süha Süreyya Özbek

PURPOSE

High-risk breast lesions (HRLs) are associated with future risk of breast cancer. Considering the pathological subtypes, malignancy upgrade rate differs according to each subtype and depends on various factors such as clinical and radiological features and biopsy method. Using artificial intelligence and machine learning models in breast imaging, evaluations can be made in terms of risk estimation in different research areas. This study aimed to develop a machine learning model to distinguish HRL cases requiring surgical excision from lesions with a low risk of accompanying malignancy.

METHODS

A total of 94 patients who were diagnosed with HRL by image-guided biopsy between January 2008 and March 2020 were included in the study. A structured database was created with clinical and radiological characteristics and histopathological results. A machine learning prediction model was created to make binary classifications of lesions as malignant or benign. Random forest, decision tree, K-nearest neighbors, logistic regression, support vector machine (SVM), and multilayer perceptron machine learning algorithms were used. Among these algorithms, SVM was the most successful. The estimations of malignancy for each case detected by artificial intelligence were combined and statistical analyses were performed.

RESULTS

Considering all cases, the malignancy upgrade rate was 24.5%. A significant association was observed between malignancy upgrade rate and lesion size ($P = 0.004$), presence of mammography findings ($P = 0.022$), and breast imaging-reporting and data system category ($P = 0.001$). A statistically significant association was also found between the artificial intelligence prediction model and malignancy upgrade rate ($P < 0.001$). With the SVM model, an 84% accuracy and 0.786 area-under-the-curve score were obtained in classifying the data as benign or malignant.

CONCLUSION

Our artificial intelligence model (SVM) can predict HRLs that can be followed up with a lower risk of accompanying malignancy. Unnecessary surgeries can be reduced, or second line vacuum excisions can be performed in HRLs, which are mostly benign, by evaluating on a case-by-case basis, in line with radiology–pathology compatibility and by using an artificial intelligence model.

KEYWORDS

Artificial intelligence, breast, cancer, high risk lesion of breast, image-guided biopsy

From the Department of Radiology (Ö.A. ✉ dr.ozgeaslan@gmail.com, A.O., S.S.Ö.), Ege University Faculty of Medicine, İzmir, Turkey; Department of Computer Engineering (B.K., R.C.E., O.D.), Ege University Engineering Faculty, İzmir, Turkey; Department of General Surgery (L.Y.), Ege University Faculty of Medicine, İzmir, Turkey; Department of Medical Pathology (O.Z.), Ege University Faculty of Medicine, İzmir, Turkey.

Received 21 November 2021; revision requested 20 December 2021; last revision received 13 March 2022; accepted 16 March 2022.

Epub: 24.01.2023

Publication date: 29.03.2023

DOI: 10.5152/dir.2022.211047

The increase in breast cancer screening with mammography increases the rate of non-palpable lesions detected in the breast.[1,2] In the diagnosis of these lesions, percutaneous biopsy methods are increasingly applied under the guidance of imaging methods. Percutaneous needle biopsy is a fast, easy-to-apply, inexpensive, and well-tolerated biopsy alternative to open surgical biopsies.[1,3] The prevalence of high-risk breast lesion (HRL) detection with core needle biopsy (CNB) is 5–9% in all breast biopsies.[2,4,5] HRLs are defined as lesions with a high risk of malignant transformation and the possibility of synchronous or ad-

You may cite this article as: Aslan Ö, Oktay A, Katuk B, et al. Prediction of malignancy upgrade rate in high-risk breast lesions by an artificial intelligence model: a retrospective study. *Diagn Interv Radiol.* 2023;29(2):260-267.

Table 1. Benign and malignant lesion distribution characteristics

		Benign group	Malignant group	SVM model prediction result (n)		Total
		n (%)	n (%)	Wrong	Right	n (%)
Biopsy type	CNB	55 (59)	15 (16)	8	62	70 (74)
	VAB	16 (17)	8 (9)	1	23	24 (26)
Risk	Positive	25 (27)	8 (9)	1	32	33 (35)
	Negative	46 (49)	15 (16)	8	53	61 (65)
Needle thickness	9 G	16 (17)	8 (9)	1	23	24 (26)
	14 G	55 (59)	15 (16)	8	62	70 (74)
Sampling number	<4	15 (16)	4 (4)	3	16	19 (20)
	≥4	56 (60)	19 (20)	6	69	75 (80)
Lesion diameter	<1.5 cm	46 (49)	7 (7)	4	49	53 (56)
	≥1.5 cm	25 (27)	16 (17)	5	36	41 (44)
Mammography finding	Microcalcification	19 (40)	10 (21)	1	28	29 (62)
	Mass	6 (13)	3 (6)	1	8	9 (19)
	Non-mass	6 (13)	3 (6)	2	7	9 (19)
US finding	Mass	40 (56)	10 (14)	4	46	50 (70)
	Non-mass	14 (20)	7 (10)	4	17	21 (30)
MRI finding	Mass	16 (41)	3 (8)	1	18	19 (49)
	Non-mass	13 (33)	7 (18)	5	15	20 (51)
Microcalcification morphology	Amorphous, course heterogeneous	11 (38)	4 (14)	1	14	15 (52)
	Fine linear branching, fine pleomorphic	8 (28)	6 (21)	-	14	14 (48)

CNB, core needle biopsy; VAB, vacuum assisted biopsy; SVM, support vector machine; US, ultrasonography; MRI, magnetic resonance imaging.

Table 2. Distribution characteristics of pathological subgroups and malignancy upgrade rates

Pathologic subgroups	Benign group n (%)	Malignant group n (%)	Total n (%)	Malignancy upgrade rate (%)
IP	31 (33)	4 (4)	35 (37)	11.4
ADH	29 (31)	13 (14)	42 (45)	31
AIP	7 (7)	3 (3)	10 (11)	30
Radial scar	3 (3)	2 (2)	5 (5)	40
Lobular neoplasia	1 (1)	1 (1)	2 (2)	50
Total	71 (76)	23 (24)	94 (100)	24.5

IP, intraductal papilloma without atypia; ADH, atypical ductal hyperplasia; AIP, atypical intraductal papilloma.

Table 3. Variables that have a statistically significant relationship with malignancy upgrade rate

		Benign group n (%)	Malignant group n (%)	Total n (%)	*P* value
BI-RADS category	3-4A-4B	64 (68)	14 (15)	78 (83)	0.003
	4C-5	7 (7)	9 (10)	16 (17)	
Lesion diameter	<1.5 cm	46 (49)	7 (7)	53 (56)	0.004
	≥1.5 cm	25 (27)	16 (17)	41 (44)	
Mammographic finding	Positive	33 (35)	17 (18)	50 (53)	0.022
	Negative	38 (40)	6 (6)	44 (47)	

BI-RADS, breast imaging reporting and data system.

The AI analysis identified 85 cases correctly and 9 cases incorrectly (Tables 1 and 4). The SVM AI model, which was trained using certain hyperparameters, had 84% accuracy (Figure 5) and an AUC score of 0.786 (Figure 7) in classifying the data as benign or malignant.

No statistically significant difference was found between needle thickness/biopsy type and erroneous AI estimation ($P = 0.297$).

A statistically significant difference was found between the AI prediction and the malignancy upgrade rate of the patients ($P < 0.001$). The sensitivity of the malignant case prediction set of the AI model was 60.87%, the specificity was 100%, PPV was 100%, and negative predictive value was 88.75%.

Discussion

The most significant problem in the management of HRLs is upgrading to malignancy. The upgrade rate to malignancy in this study was 24.5%, which is similar to the rates reported in the literature.[10]

Considering pathological subtypes, the rate of upgrade to malignancy differs according to each subtype. The malignancy upgrade rate of ADH, which was the most common lesion subtype among our cases, was similar to the literature. A wide range of malignancy upgrade rates for ADH and AIP has been reported in the literature.[2,7] In this study, for AIP, as in ADH, there were erroneous AI predictions in three of the patients,

Table 4. Distribution characteristics of pathological subgroups malignancy upgrade rates and the accuracy of the SVM model in subtypes

Pathologic subgroup	True malignancy upgrade rates n (%)	SVM model malignancy upgrade rates n (%)	SVM model accuracy (%)
ADH (n = 42)	13 (31)	10 (23.8)	92.8
AIP (n = 10)	3 (30)	-	70
LN (n = 2)	1 (50)	1 (50)	100
RS (n = 5)	2 (40)	-	60
IP (n = 35)	4 (11.4)	3 (8.6)	97.1
Total	23 (24.5)	14 (14.9)	90.4

ADH, atypical ductal hyperplasia; AIP, atypical intraductal papilloma; LN, lobular neoplasia; RS, radial scar; IP, intraductal papilloma without atypia; SVM, support vector machine.

and biopsies were performed with tru-cut in both groups.

The malignancy upgrade rates for radial scar and lobular neoplasia were in the upper limit of the rates stated in the literature.[7] This may be due to the low number of cases in these subgroups.

The SVM model made an incorrect prediction in nine malignant cases in total (Tables 1 and 4). One of these cases was diagnosed by VAB with a 9-G needle, and all others were diagnosed by CNB with a 14-G needle. No statistically significant correlation was found between needle thickness/biopsy type and erroneous AI estimation, but the low number of cases is a limitation in the evaluation of this variable. In the study of Bahl et al.[13], which included 1,006 HRLs, the AI prediction model had a prediction accuracy of 97.4% in malignant cases and 69.4% in benign cases, and they reported that unnecessary surgeries could be reduced in benign cases. In the present study, the AI model made a correct prediction in all cases that were diagnosed as benign by surgical excision and considered as stable in long-term follow-up. More than half of the patients who underwent surgical excision were diagnosed as benign. Considering the radiopathological fit and AI model estimation, if these cases had been followed up radiologically and clinically, the rate of unnecessary surgery could have been reduced by 71%.

The majority of HRLs are benign but most are surgically excised because of the associated risk of malignancy. Post-biopsy evaluation and biopsy procedure are important in the management of these lesions.[1,15]

Comparable to the literature, there was a statistically significant relationship between VAB as a biopsy guide method, 9-G needle thickness, and sufficient number of samples with the malignancy upgrade rate. In cases that have these features, a more appropriate decision can be made in terms of follow-up and excision. In order to increase the correct prediction rates with the AI model, studies containing more cases and data sets are needed.

There are some further limitations to this study. Firstly, it is a retrospective study. The differences in the number of pathological subtypes and the low number of patients were our biggest limitations. Significant results could not be obtained in many statistical analyses due to the differences in the number of pathological subtypes such as lobular neoplasia and radial scar, and the low number of cases. In addition, due to the small number of cases and the limited number of histopathological features in terms of the degree of atypia, a clear analysis of the variables that may be effective in the erroneous predictions of the SVM model could not be made. For this reason, better statistical results can be obtained by adding features such as the degree of pathological atypia, which will further strengthen the data set, and by including more patients.

In conclusion, this study's AI model (SVM) can predict HRLs that can be followed up with a lower risk of accompanying malignancy. Both ADH and AIP cases should be surgically excised because of the high risk of malignancy associated with them. Apart from these subtypes, HRLs, which are mostly benign, can be evaluated on a case-by-case basis, in line with radiology–pathology compatibility and using an AI prediction model, to reduce unnecessary surgeries, or excision can be performed with second-line VAB.

Conflict of interest disclosure

The authors declare no conflicts of interest.

References

1. Oktay A. Meme hastalıklarında görüntüleme. İçinde: Tunacı M, ed. Radyoloji-patoloji sonuçlarının korelasyonu, bölüm 7. 2nd ed. Ankara: *Dünya Tıp Kitabevi*; 2020:249-255. [Crossref]
2. Neal L, Sandhu NP, Hieken TJ, et al. Diagnosis and management of benign, atypical, and indeterminate breast lesions detected on core needle biopsy. *Mayo Clin Proc.* 2014;89(4):536-547. [Crossref]
3. Liberman L. Clinical management issues in percutaneous core breast biopsy. *Radiol Clin North Am.* 2000;38(4):791-807. [Crossref]
4. Oktay A. Radyolojik-patolojik korelasyon: yüksek risk lezyonlarda ne yapmalıyız? *Türk Radyoloji Seminerleri.* 2014;2:217-229. [Crossref]
5. Gulla S, Lancaster R, De Los Santos J. High-risk breast lesions and current management. *Semin Roentgenol.* 2018;53(4):252-260. [Crossref]
6. Bianchi S, Caini S, Renne G, et al. Positive predictive value for malignancy on surgical excision of breast lesions of uncertain malignant potential (B3) diagnosed by stereotactic vacuum-assisted needle core biopsy (VANCB): a large multi-institutional study in Italy. *Breast.* 2011;20(3):264-270. [Crossref]
7. Nakhlis F. How do we approach benign proliferative lesions? *Curr Oncol Rep.* 2018;20(4):34. [Crossref]
8. Georgian-Smith D, Lawton TJ. Controversies on the management of high-risk lesions at core biopsy from a radiology/pathology perspective. *Radiol Clin North Am.* 2010;48(5):999-1012. [Crossref]
9. Thomas PS. Diagnosis and management of high-risk breast lesions. *J Natl Compr Canc Netw*. 2018;16(11):1391-1396. [Crossref]
10. Boateng S, Tirada N, Khorjekar G, Richards S, Ioffe O. Excision or observation: the dilemma of managing high-risk breast lesions. *Curr Probl Diagn Radiol.* 2020;49:124-132. [Crossref]
11. Rageth CJ, O'Flynn EAM, Pinker K, et al. Second international consensus conference on lesions of uncertain malignant potential in the breast (B3 lesions). *Breast Cancer Res Treat.* 2019;174(2):279-296. [Crossref]
12. Mendelson EB. Artificial intelligence in breast imaging: potentials and limitations. *AJR Am J Roentgenol.* 2019;212(2):293-299. [Crossref]
13. Bahl M, Barzilay R, Yedidia AB, Locascio NJ, Yu L, Lehman CD. High-risk breast lesions: a machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision. *Radiology.* 2018;286(3):810-818. [Crossref]
14. Murtaza G, Shuib L, Wahab AWA, et al. Deep learning-based breast cancer classification through medical imaging modalities: state of

the art and research challenges. *Artif Intell Rev.* 2019;53:1655-1720. [Crossref]

15. Krishnamurthy S, Bevers T, Kuerer H, Yang WT. Multidisciplinary considerations in the management of high-risk breast lesions. *AJR Am J Roentgenol.* 2012;198(2):132-140. [Crossref]
16. Armes JE, Galbraith C, Gray J, Taylor K. The outcome of papillary lesions of the breast diagnosed by standard core needle biopsy within a BreastScreen Australia service. *Pathology.* 2017;49(3):267-270. [Crossref]