

A Pre-operative Score to Discriminate Fibroepithelial Lesions of the Breast: Phyllode Tumor or Fibroadenoma?

CAMILLE MIMOUN^{1,2,3}, AMELIE ZELLER², JULIEN SEROR¹, LAURENE MAJOULET¹,
EVA MARCHAND², MATTHIEU MEZZADRI², PASCAL CHEREL¹, VINCIANE PLACE²,
FRANÇOISE CORNELIS², JEAN-LOUIS BENIFLA² and ROMAN ROUZIER¹

¹Department of Surgical Oncology, Rene Huguenin, Saint-Cloud, France;

²Department of Gynecology and Obstetrics, Lariboisiere Hospital, Paris, France;

³Research unit EA 7285 "Risk and safety in clinical medicine for women and perinatal health",
University of Versailles Saint-Quentin (UVSQ), Versailles, France

Abstract. *Background/Aim:* Fibroepithelial lesions (FEL) of the breast include fibroadenomas and phyllodes tumors (PT). Their histologic characteristics on core needle biopsy can overlap, while their clinical management is different. The aim of this study was to develop and to validate a pre-operative score for the diagnosis of PT with surgical decision rules. *Patients and Methods:* We developed a pre-operative score for the diagnosis of PT by performing logistic regression on 217 FEL of the Rene Huguenin Hospital. This score and the surgical decision rules were validated on 87 FEL of the Lariboisiere Hospital. *Results:* Three variables were independently and significantly associated with PT: age ≥ 40 years, mammography's tumor size ≥ 3 cm and PT diagnosed by CNB. The pre-operative score was based on these three criteria with values ranging from 0 to 10. Surgical decision rules were created: the low-risk group of PT (score ≤ 2) had a sensitivity of 92.6% and a LR- of 0.2, the high-risk group (score > 7) had a specificity of 93.5% and a LR+ of 4.4. In the validation sample, surgical decision rules were applied. *Conclusion:* These surgical decision rules may prove useful in deciding which FEL needs surgical resection.

Fibroepithelial lesions (FEL) of the breast are heterogeneous biphasic neoplasms involving proliferation of both epithelial and stromal components including the common benign fibroadenomas (FA) and the rarer phyllodes tumors (PT) (1-3). In 2012, the World Health Organization (WHO) sub-classified PT histologically as benign, borderline or malignant grade categories based on the degree of stromal cellularity and atypia, mitotic count, stromal overgrowth and tumor borders (1).

Correspondence to: Camille Mimoun, 2 Rue Ambroise Paré, 75010, Paris, France. E-mail: mimouncamille@gmail.com

Key Words: Fibroepithelial lesions, phyllodes tumors, fibroadenoma, score.

PT and FA have different clinical implications and surgical management. Whereas PT can be aggressive with possibility of recurrence, potential for malignant transformation and distant metastasis, FA can regress with age (4). A complete surgical excision with an attempt to obtain negative margins of 1 cm or greater is the current National Comprehensive Cancer Network (NCCN) guideline recommendation for PT. The management of FA can be follow-up evaluation and reassurance (5).

Differences in the recommended management of these lesions, ranging from observation to wide surgical resection, make the distinction between PT and FA crucial. However, FEL can be difficult to diagnose with an initial core needle biopsy (CNB). A CNB demonstrating FEL with any histologic feature suggesting a more advanced lesion than FA may indicate a PT (1-3, 6, 7). Consequently, surgical excision is recommended for many patients to rule out PT resulting in potential overtreatment (1-3, 8, 9).

The aim of this study was to develop and to validate a pre-operative score for the diagnosis of PT with surgical decision rules.

Patients and Methods

Study design and database. We conducted a bicenter retrospective study in FEL of the breast diagnosed by CNB at two hospitals. FEL from the Department of Surgical Oncology of the Rene Huguenin Hospital (Saint-Cloud, France) were included from April 1997 to November 2015 and those from the Department of Gynecology and Obstetrics of the Lariboisiere Hospital were included from October 2006 to November 2016. Our score was developed using data from the population of the Rene Huguenin Hospital and validated on the data from patients of the Lariboisiere Hospital.

We included all operated lesions. Exclusion criteria were recurrences of PT, lesions belonging to patients lost to follow-up after CNB or with history of breast cancer.

All CNBs were performed by breast specialized radiologists. All CNBs and surgical specimens were reviewed by a breast specialized pathologist. All surgical excisions were made by a breast specialized surgeon.

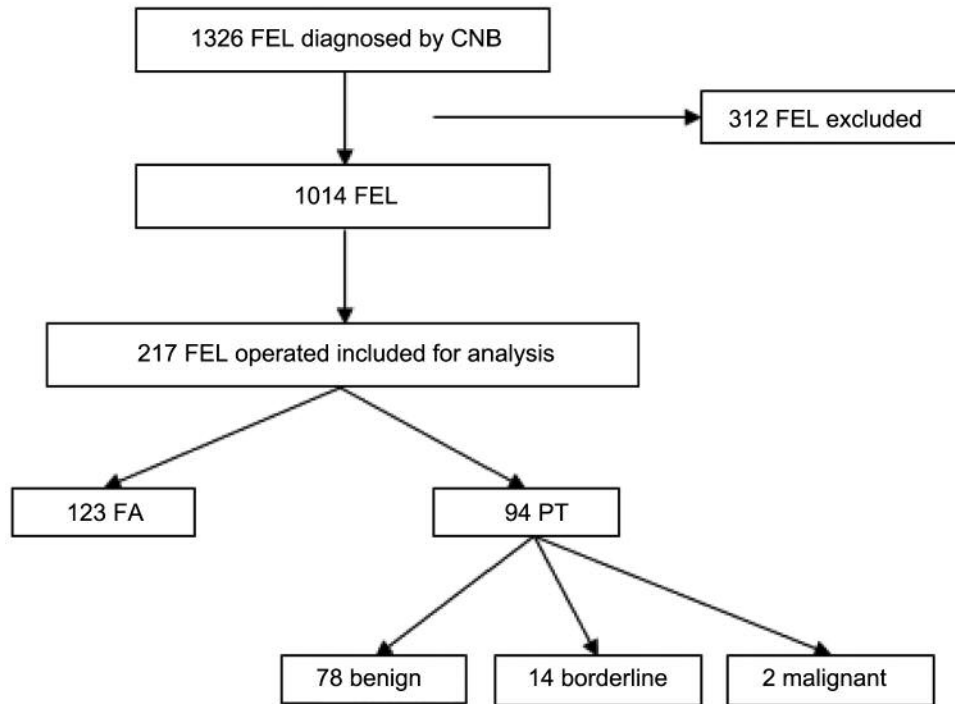


Figure 1. Flow chart for the derivation sample.

Our database consisted of clinical, radiological, surgical and histological information.

Diagnosis of FEL. The diagnostic reference (gold standard) used for the diagnosis of FEL, FA or PT, was the final pathologic diagnosis on the surgical specimen. PT were described with the 2012 WHO classification as benign, borderline or malignant (1).

Statistical analysis. We compared FEL defined as FA to those defined as PT.

We carried out univariate analysis using a quantitative (Student’s *t*-test) or qualitative (Chi² test) test as appropriate. We converted variables associated with the presence of PT at a threshold of *p*<0.10 into dichotomous variables using receiver operating characteristic (ROC) curves.

We used multiple logistic regression analysis to select the best combination of variables that was independently associated with the diagnosis of PT (*p*<0.05). Variables were selected by a backward stepwise procedure from those associated with PT in the univariate analysis at a threshold of *p*<0.10.

Bootstrap resampling was performed to assess the robustness of the multivariate model, using 1,000 replications, to estimate the distribution of each logistic regression coefficient to remove variables potentially responsible for instability in the model (10).

The performance of the model in the diagnosis of PT was specified by calculating its area under the ROC curve (ROC-AUC).

The logistic regression model was used to build a score by rounding up the β coefficients from the multivariate analysis to generate a simple scale (11). Missing data were considered as absent. The ROC-AUCs of the logistic regression model and score

model were compared to check that the two values were not significantly different. Sensibility, specificity, positive and negative likelihood ratios were calculated for different threshold values of the score.

We made surgical decision rules by classifying patients as at low-risk or high-risk of PT by choosing two threshold values of the score to make two classifications, one with sensitivity >95% and positive likelihood ratio (LR-) <0.25 (rule out) and the other with specificity >90% and negative likelihood ratio (LR+) >4.0 (rule in) (12).

Finally, the surgical decision rules were applied to the validation sample for external validation, by calculating sensitivity, specificity, LR+ and LR- with their 95%CI in the low-risk and the high-risk groups of PT.

Statistical analyses were performed using stat 13.0 (Stata Corp; College Station, TX, USA).

Results

FEL. A total of 1326 FEL were diagnosed by CNB in the derivation sample. Among them, 312 were excluded (4 recurrences of PT, 291 lesions belonging to patients lost to follow-up after CNB and 17 lesions belonging to patients with a history of breast cancer). Of the 1014 FEL left, 217 benefited surgical excision. There were 123 FA and 94 PT (78 benign, 14 borderline and 2 malignant) (Figure 1).

The validation sample consisted of 87 FEL diagnosed by CNB and operated on. There were 68 FA and 19 PT (16 benign, 2 borderline and 1 malignant).

Table I. Univariate analysis of the derivation sample.

Variables	FA N=123 Mean+/-SD ou n (%) ^a	PT N=94 Mean+/-SD ou n (%) ^a	p-Value
Age (years)	38.8+/-10.6	43.4+/-12.2	0.004 ^b
Tumor size (TNM classification)			0.121 ^c
T0	60 (48.8)	30 (31.9)	
T1	30 (24.4)	26 (27.7)	
T2	27 (22.0)	29 (30.9)	
T3	5 (4.1)	8 (8.5)	
T4	1 (0.8)	1 (1.1)	
Mammography's tumor size (mm)	23.1+/-12.5	31.9+/-29.0	0.021 ^b
Sonography's tumor size (mm)	21.4+/-13.4	29.2+/-26.0	<0.001 ^b
Mammography's ACR			0.01 ^c
1	28 (23.7)	14 (16.3)	
2	1 (0.9)	0 (0.0)	
3	33 (28.0)	11 (12.8)	
4	54 (45.8)	57 (66.3)	
5	2 (1.7)	4 (4.7)	
Sonography's ACR			0.01 ^c
1	2 (1.8)	0 (0.0)	
2	0 (0.0)	0 (0.0)	
3	45 (39.5)	29 (31.5)	
4	66 (57.9)	56 (60.9)	
5	1 (0.9)	7 (7.6)	
ACR max			0.04 ^c
1	0 (0.0)	0 (0.0)	
2	0 (0.0)	0 (0.0)	
3	35 (28.5)	17 (18.1)	
4	85 (69.1)	69 (73.4)	
5	3 (2.4)	8 (8.5)	
Histological result of the CNB			<0.001 ^c
FA	58 (47.2)	10 (10.6)	
PT	65 (52.9)	84 (89.4)	

^aPercentages are calculated with the total number of FEL despite of some missing data; ^bChi2 test; ^cStudent's *t*-test. ACR: American Classification of Radiology; SD: standard deviation; TNM: tumor, nodes, metastases.

Univariate and multivariate analyses. The findings from the univariate analysis are reported in Table I.

The multiple logistic regression analysis identified three variables independently and significantly ($p < 0.05$) associated with the diagnosis of PT, namely, age ≥ 40 years ($p = 0.026$), mammography's tumor size ≥ 30 mm ($p < 0.001$) and PT diagnosed by CNB ($p < 0.001$) (Table II). All of these variables were stable after 1000 bootstrap replications (Table II).

The ROC-AUC of this model was 0.76 95%CI=0.74-0.86 (Figure 2).

Table II. Multiple logistic regression model.

Variables	p-Value	p-Value Bootstrap
Age	0.026	0.035
<40 years		
≥ 40 years		
Mammography's tumor size	0.002	0.003
<30 mm		
≥ 30 mm		
Histological result of the CNB	<0.001	<0.001
FA		
PT		

CNB: Core needle biopsy; FA: fibroadenomas; PT: phyllodes tumors.

Score and surgical decision rules. The score was given by the following equation: score=(age ≥ 40 years $\times 2$)+(mammography's tumor size ≥ 3 cm $\times 3$)+(PT diagnosed by CNB $\times 5$) (Table III).

The ROC-AUC of the score was 0.75 (95%CI=0.69-0.82) (Figure 3). There was no significant difference between the ROC-AUC of the score and the ROC-AUC of the model ($p = 0.78$).

We then defined surgical decision rules with a low and high-risk group of PT (Table IV): FEL with a score ≤ 2 were defined as a low-risk group of PT. A threshold score value of 2 produced a sensitivity of 92.6% (95%CI=85.3-97.0) and a LR- of 0.2 (95%CI=0.1-0.4). A surgical excision is not needed for those FEL; FEL with a score > 7 were defined as a high-risk group of PT. A threshold value of 7 produced a specificity of 93.5% (95%CI=87.6-97.2) and a LR+ of 4.4 (95%CI=2.1-9.3). A surgical excision is absolutely needed for those FEL with 1 cm margins.

The surgical decision rules were applied in the validation sample. The low-risk group of PT had a sensitivity of 84.2% (95%CI=60.4-96.6) and a LR- of 0.24 (95%CI=0.09-0.7). The high-risk group of PT had a specificity of 98.5% (95%CI=92.1-100.0) and a LR+ of 17.9 (95%CI=2.2-144.0). The 95%CI of the sensitivity, the specificity, the LR+ and the LR- in the validation sample were in the expected range.

Discussion

We have developed the first pre-operative score for the diagnosis of PT with surgical decision rules for FEL diagnosed by CNB of the breast. The 10-point score is based on three simple items: age ≥ 40 years (2 points), mammography's tumor size ≥ 3 cm (3 points) and PT diagnosed by CNB (5 points). The surgical decision rules were obtained by classifying FEL as at low-risk or high-risk of PT. A score ≤ 2 puts FEL in a low-risk group of PT with a sensitivity of 92.6 and a LR- of 0.2, and these FEL can be observed; a score > 7 puts patient in a high-risk group of PT

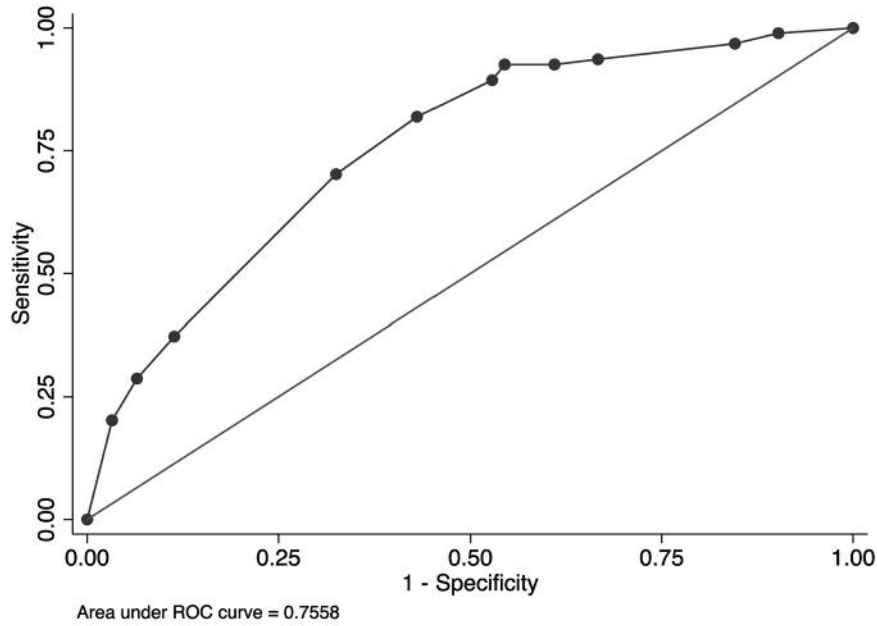


Figure 2. ROC curve of the model.

with a specificity of 93.5% and a LR+ of 4.4, and these FEL need wide resection with 1 cm margins. These surgical decision rules were validated in an external sample.

The strengths of our study are the following. The first is the face validity of our score. Indeed, the score is made of one clinical, one imaging and one histological item. Furthermore, these three items are validated in the literature to be helpful in distinguishing FA from PT. Some studies have demonstrated that the median age at presentation of PT is 42-45 years (13) and that FA have a peak of incidence at 20-40 years (14). Also, a diagnosis of PT may be favored if the tumor is larger than 4 cm (15, 16). The goals of the triple approach, clinical, imaging and histological evaluation, are to find all PT among FEL and so avoid unnecessary diagnostic surgical procedures, if possible. However, it has been suggested that even the use of this approach may lead to disappointing results (16) because clinical, imaging and histological characteristics of these lesions can overlap (1-3, 6, 17). Our model is the first to reunite these pre-operative characteristics to obtain better performance in the pre-operative diagnosis of PT than each item considered separately.

Second is that we took into account overfitting, the most important bias in predictive studies (18). We used two statistical methods to minimize this risk: the bootstrap resampling procedure, which allowed us to detect and remove variables potentially responsible for instability in the model (10) and the validation in a specific population (19).

Table III. Score.

Variables	Score points
Age	
<40 years	0
≥40 years	2
Mammography's tumor size	
<30 mm	0
≥30 mm	3
Histological result of the CNB	
FA	0
PT	5
Total	/10 points

CNB: Core needle biopsy; FA: fibroadenomas; PT: phyllodes tumors.

Third, although the study was retrospective, as the data were not initially gathered for this study, the database of the derivation sample was prospectively constituted from all FEL of the breast diagnosed by CNB since 1997 in the Department of Surgical Oncology of the Rene Huguenin Hospital. This ensured that all operated FEL were included in our study to build the score.

The principal limitation of our study is the referral bias that could have occurred. Indeed, we only included FEL which underwent surgical resection implying that observed lesions were not included. Those FEL may differ from those which underwent surgical resection: it can be assumed that

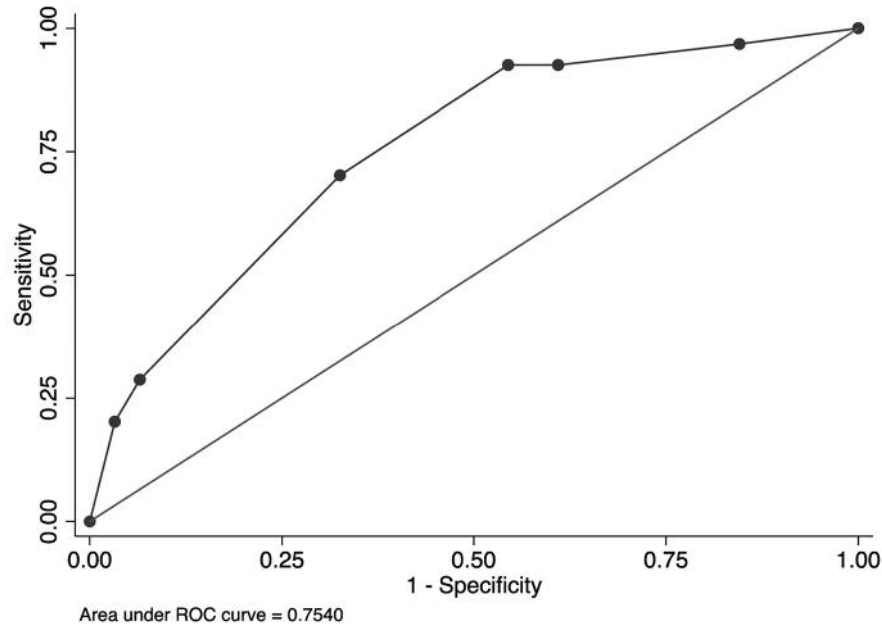


Figure 3. ROC curve of the score.

operated FEL have a higher pre-operative suspicion of PT than observed lesions as attested by the high prevalence of PT in our population. This could have increased the link between the variables studied and PT and may have affected the diagnostic accuracy of our model (20).

We have developed a 10-point score which is reproducible, original, and would be easy to use in routine practice. In daily clinical practice, our pre-operative score could be helpful to a multidisciplinary team to plan the surgical management of patients with FEL on the CNB of the breast. Patients belonging to the low-risk PT group would not be eligible for surgical resection, whereas those belonging to the high-risk group should undergo wide surgical resection with 1 cm margins.

Conflicts of Interest

The Authors have no conflicts of interest to declare regarding this study.

Authors' Contributions

Camille Mimoun and Roman Rouzier contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Amelie Zeller, Julien Seror, Laurene Majoulet, Eva Marchand, Matthieu Mezzadri, Pascal Cherel, Vinciane Place, Françoise Cornelis and Jean-Louis Benifla contributed to the design and implementation of the research.

Table IV. Diagnostic values of the score for the low and high-risk groups in the derivation sample.

Score	Sensitivity % 95%CI	Specificity % 95%CI	LR+ 95%CI	LR- 95%CI
≤2	92.6* 85.3-97.0	39.0 30.4-48.2	1.5 1.3-1.7	0.2 0.1-0.4
>7	28.7 19.9-39.0	93.5 87.6-97.2	4.4 2.1-9.3	0.7 0.7-0.9

*Calculated for the absence of PT. LR: Likelihood ratio.

References

- Lakhani SR, Ellis IO, Schnitt SJ, Tan PH and van de Vijver MJ: WHO Classification of Tumours of the Breast. 4th Edition, Volume 4. Available from: <http://publications.iarc.fr/Book-And-Report-Series/Who-Iarc-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012> (Last accessed 19/12/19)
- Tan PH and Ellis IO: Myoepithelial and epithelial-myoepithelial, mesenchymal and fibroepithelial breast lesions: updates from the WHO Classification of Tumours of the Breast 2012. *J Clin Pathol* 66(6): 465-470, 2013. PMID: 23533258. DOI: 10.1136/jclinpath-2012-201078
- Karim RZ, O'Toole SA, Scolyer RA, Cooper CL, Chan B, Selinger C, Yu B, Carmalt H, Mak C, Tse GM, Tan PH, Putti TC

- and Lee CS: Recent insights into the molecular pathogenesis of mammary phyllodes tumours. *J Clin Pathol* 66(6): 496-505, 2013. PMID: 23404800. DOI: 10.1136/jclinpath-2012-201082
- 4 Krings G, Bean GR and Chen YY: Fibroepithelial lesions; The WHO spectrum. *Semin Diagn Pathol* 34(5): 438-452, 2017. PMID: 28688536. DOI: 10.1053/j.semdp.2017.05.006
 - 5 Amin AL, Purdy AC, Mattingly JD, Kong AL and Termuhlen PM: Benign breast disease. *Surg Clin North Am* 93(2): 299-308, 2013. PMID: 23464687. DOI: 10.1016/j.suc.2013.01.001
 - 6 Resetkova E, Khazai L, Albarracin CT and Arribas E: Clinical and radiologic data and core needle biopsy findings should dictate management of cellular fibroepithelial tumors of the breast. *Breast J* 16(6): 573-580, 2010. PMID: 21070433. DOI: 10.1111/j.1524-4741.2010.01013.x
 - 7 Johnson NB and Collins LC: Update on percutaneous needle biopsy of nonmalignant breast lesions. *Adv Anat Pathol* 16(4): 183-195, 2009. PMID: 19546607. DOI: 10.1097/PAP.0b013e3181a9d33e
 - 8 Corben AD, Edelweiss M and Brogi E: Challenges in the interpretation of breast core biopsies. *Breast J* 16: S5-9, 2010. PMID: 21050312. DOI: 10.1111/j.1524-4741.2010.00993.x
 - 9 Jacobs TW, Connolly JL and Schnitt SJ: Nonmalignant lesions in breast core needle biopsies: to excise or not to excise? *Am J Surg Pathol* 26(9): 1095-1110, 2002. PMID: 12218567. DOI: 10.1097/00000478-200209000-00001
 - 10 Efron B and Gong G: A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am Stat* 37(1): 36-48, 1983. DOI: 10.2307/2685844
 - 11 Buckley RG, King KJ, Disney JD, Ambroz PK, Gorman JD and Klausen JH: Derivation of a clinical prediction model for the emergency department diagnosis of ectopic pregnancy. *Acad Emerg Med Off J Soc Acad Emerg Med* 5(10): 951-960, 1998. PMID: 9862584. DOI: 10.1111/j.1553-2712.1998.tb02770.x
 - 12 Hubbard JL, Cagle K, Davis JW, Kaups KL and Kodama M: Criteria for excision of suspected fibroadenomas of the breast. *Am J Surg* 209(2): 297-301, 2015. PMID: 25682095. DOI: 10.1016/j.amjsurg.2013.12.037
 - 13 Barrio AV, Clark BD, Goldberg JI, Hoque LW, Bernik SF and Flynn LW: Clinicopathologic features and long-term outcomes of 293 phyllodes tumors of the breast. *Ann Surg Oncol* 14(10): 2961-2970, 2007. PMID: 17562113. DOI: 10.1245/s10434-007-9439-z
 - 14 Ilvan S: Fibroepithelial tumors of the breast. *In: Breast disease: Diagnosis and pathology.* Aydiner A, Igci A, Soran A (eds.). Cham, Switzerland, Springer International Publishing, pp. 235-239, 2016. DOI: 10.1007/978-3-030-04606-4_16
 - 15 Hoda SA, Brogi E, Koerner FC and Rosen PP: Rosen's breast pathology. Lippincott Williams & Wilkins, 2001.
 - 16 Foxcroft LM, Evans EB and Porter AJ: Difficulties in the pre-operative diagnosis of phyllodes tumours of the breast: a study of 84 cases. *Breast Edinb Scotl* 16(1): 27-37, 2007. PMID: 16876413. DOI: 10.1016/j.breast.2006.05.004
 - 17 Simsir A and Cangiarella J: Challenging breast lesions: pitfalls and limitations of fine-needle aspiration and the role of core biopsy in specific lesions. *Diagn Cytopathol* 40(3): 262-722, 2012. PMID: 21309013. DOI: 10.1002/dc.21630
 - 18 Altman DG and Royston P: What do we mean by validating a prognostic model? *Stat Med* 19(4): 453-473, 2000. PMID: 10694730. DOI: 10.1002/(sici)1097-0258(20000229)19:4<453::aid-sim350>3.0.co;2-5
 - 19 Steyerberg E: Clinical prediction models: A practical approach to development, validation, and updating. Available from: <https://www.springer.com/gp/book/9780387772431>
 - 20 Miller TD, Hodge DO, Christian TF, Milavetz JJ, Bailey KR and Gibbons RJ: Effects of adjustment for referral bias on the sensitivity and specificity of single photon emission computed tomography for the diagnosis of coronary artery disease. *Am J Med* 112(4): 290-297, 2002. PMID: 11893368. DOI: 10.1016/s0002-9343(01)01111-1

Received December 14, 2019

Revised December 19, 2019

Accepted December 30, 2019