HAL
open science

Prospective Multicenter Study Validate a Prediction Model for Surgery Uptake Among Women with Atypical Breast Lesions

Catherine Uzan, Chafika Mazouni, Caroline Rossoni, Brigitte de Korvin, Christine Tunon de Lara, Monique Cohen, Nathalie Chabbert, Sonia Zilberman, Veronique Boussion, Anne Vincent Salomon, et al.

To cite this version:

Catherine Uzan, Chafika Mazouni, Caroline Rossoni, Brigitte de Korvin, Christine Tunon de Lara, et al.. Prospective Multicenter Study Validate a Prediction Model for Surgery Uptake Among Women with Atypical Breast Lesions. Annals of Surgical Oncology, 2020, 10.1245/s10434-020-09107-z . hal-03013819

HAL Id: hal-03013819

https://hal.sorbonne-universite.fr/hal-03013819

Submitted on 19 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prospective multicenter study to validate a prediction model for surgery uptake among women with atypical breast lesions

Running head: Surgery uptake in atypical breast lesions

Key words: Atypical breast lesion, B3 lesion, breast cancer, model, biopsy, Ductal carcinoma in situ, atypical ductal hyperplasia

Catherine Uzan MD, PhD [1], Chafika Mazouni MD, PhD [2], Caroline Rossoni MSc [2], Brigitte De Korvin MD [3], Christine Tunon de Lara MD [4], Monique Cohen MD [5], Nathalie Chabbert MD PhD[6], Sonia Zilberman MD[6], Veronique Boussion MD[7], Anne Vincent Salomon MD PhD [8], Marc Espie MD [9], Charles Coutant MD PhD [10], Frederic Marchal MD PhD [11], Flore Salviat MD [12], Loic Boulanger MD [13], Isabelle Doutriaux-Dumoulin MD [14], Eva Jouve MD [15], Carole Mathelin MD PhD [16], Pierre de Saint hilaire MD [17], Joelle Mollard [18], Corinne Balleyguier MD PhD[2], Natacha Joyon MD PhD[2], Magali Lacroix Triki MD PhD[2], Suzette Delaloge MD[2], Stefan Michiels PhD[18]

[1] AP-HP (Assistance Publique des Hôpitaux de Paris), Department of Gynecological and Breast Surgery and Oncology, Pitié-Salpêtrière University Hospital, Paris, France; Sorbonne University, INSERM UMR_S_938, "Cancer Biology and Therapeutics", Centre de Recherche Saint-Antoine (CRSA), Paris, France Institut; Universitaire de Cancérologie (IUC), Paris, France

[2] Gustave Roussy, 94805 Villejuif, France;

[3] Radiology Center, Centre Eugène Marquis, CHU Rennes, Avenue de la Bataille Flandres-Dunkerque, 35000 Rennes, France

[4] Institut Bergonié, 229 Cours de l'Argonne, 33000 Bordeaux, France

5 Institut Paoli Calmettes, 232 Boulevard de Sainte-Marguerite, 13009 Marseille, France

6 Hôpital Tenon, 4 rue de la Chine, Paris, Sorbonne University, France

7 Centre Jean Perrin 58 Rue Montalembert, 63011 Clermont-Ferrand, France

8 Institut Curie, Université Paris-Sciences Lettres, INSERM U934, Département de Médecine Diagnostique et Théranostique, 26 Rue d'Ulm, 75005 Paris, France

9 University of Paris, Hopital Saint Louis, APHP, 1 Avenue Claude Vellefaux, 75010 Paris, France

10 Centre Georges François Leclerc, 1 Rue du Professeur Marion, 21000 Dijon, France

11 Institut de Cancérologie de Lorraine, 6 Avenue de Bourgogne, 54519 Vandœuvre-lès-Nancy, France

12 Service de Biostatistique et d'Épidémiologie, Gustave Roussy, 94805 Villejuif, France; CESP INSERM U1018, Université Paris-Sud, Université Paris-Saclay, 94805 Villejuif, France.

13 Centre Oscar Lambret, 3 Rue Frédéric Combemale, 59000 Lille, France

14 Institut Cancérologique de l'ouest, Boulevard Professeur Jacques Monod, 44800 Saint-Herblain, France

15 Institut Claudius Regaud-Oncopole, 1 avenue Irène Joliot-Curie 31059 TOULOUSE

16 Les Hôpitaux universitaires de Strasbourg, Hôpital de Hautepierre Avenue Molière 67098 STRASBOURG

17 Hôpital de la Croix-Rousse 103 Grande Rue de la Croix-Rousse, 69004 Lyon

18 CHU Dupuytren 2 Avenue Martin Luther King, 87000 Limoges

Corresponding author Catherine Uzan, MD, PhD

catherine.uzan@aphp.fr

AP-HP (Assistance Publique des Hôpitaux de Paris), Department of Gynecological and Breast Surgery and Oncology, Pitié-Salpêtrière University Hospital, Paris, France,

Sorbonne University, INSERM UMR_S_938, "Cancer Biology and Therapeutics", Centre de Recherche Saint-Antoine (CRSA), Paris, France

Institut Universitaire de Cancérologie (IUC), Paris, France

Synopsis

Histologically atypical breast lesions (ABLs) identified on percutaneous biopsy lead to unnecessary surgery in 75 to 90% of women.

This is the first multicenter nationwide prospective validation of a prediction model of breast cancer that allows the avoidance of unnecessary surgery among women with any ABL including atypical ductal hyperplasia (ADH).

For the pre-specified threshold of 20% predicted probability of cancer, negative predictive value was 82% (77-87%). However, at a10% probability threshold, negative predictive value was 89% (84-94%), while 58% of the patients (and 54% in patients with ADH) could have avoided surgery with only 2 missed invasive cancers.

Abstract

Background

Diagnosis of atypical breast lesions (ABLs) leads to unnecessary surgery in 75 to 90% of women. We have previously developed a model including age, complete radiological target excision after biopsy and focus size that predicts the probability of cancer at surgery. The present study aimed to validate this model in a prospective multicenter setting.

Methods

Women with a recently diagnosed ABL on image-guided biopsy were recruited in 18 centers, before wire-guided localized excisional lumpectomy. Primary outcome was the negative predictive value (NPV) of the model.

Results

The NOMAT model could be used in 287 of the 300 patients included (195 with ADH). At surgery, 12 invasive (all grade 1), and 43 in situ carcinomas were identified (all ABL: 55/287, 19%; ADH only: 49/195, 25%). The area under the receiving operating characteristics curve of the model was 0.64 (95% CI, 0.58-0.69) for all ABL; and 0.63 for ADH only (95% CI, 0.56 to 0.70)). For the pre-specified threshold of 20% predicted probability of cancer, NPV was 82% (77-87%) for all ABL and 77% (95% CI, 71% to 83%) for patients with ADH. At a 10% threshold, NPV was 89% (84-94%) for all ABL and 85% (95% CI, 78% to 92%) for the ADH. At this threshold, 58% of the whole ABL population (and 54% of ADH patients) could have avoided surgery with only 2 missed invasive cancers.

Conclusions

The NOMAT model could be useful to avoid unnecessary surgery among women with ABL, including for patients with ADH.

Clinical Trial registration: NCT02523612

Introduction

Mammographic screening, together with better treatments, has led to improved breast cancer survival over the past 30 years (1, 2). However, the currently proposed screening strategy has received some degree of criticism, especially for its associated substantial rate of overdiagnosis and overtreatment (3). Atypical breast lesions (ABL) found on percutaneous biopsies are emblematic of this issue, as they classically lead to unnecessary surgery in 75 to 90% of women (4, 5). Furthermore, in the 10-25% of women ultimately diagnosed with cancer at surgery, the latter generally has a low propensity for progression (low-grade ductal carcinoma in situ [DCIS] or invasive carcinomas). While ongoing studies (LORD NCT02492607, LORIS ISRCTN 27544579 and COMET) have been designed to determine if a simple follow-up is safe for patients with low-grade DCIS (6-8), guidelines regarding ABL lesions remain heterogeneous. The best management of these lesions (surgery or follow-up) is still a matter of debate (9).

ABLs include the following subtypes: flat epithelial atypia (FEA), atypical ductal (ADH) or lobular hyperplasia (ALH) and lobular carcinoma in situ (LCIS). They are part of breast lesions currently referred as B3 lesions (which additionally include radial scar, papillary lesions and phyllodes tumors). ABL diagnosis depends on the pathologist's experience and has limited reproducibility, especially regarding ABL subtyping. Many ABLs have mixed components, exhibiting both ADH and ALH, as well as FEA (10). Several prediction models have been proposed to determine which lesions are associated with a high risk of upgrading to cancer at surgery, and therefore may require surgical excision (11, 12). However, most of them were focused on one or another kind of ABL. Furthermore, their performances were poorly reproducible (for example AUC 0.51 (95% confidence interval: 0.47-0.53) for Ko score in an independent validation study (13)). Based on a single-institution retrospective database, we have developed a model (NOMAT, NOMogram for ATypia) capable of predicting the presence of cancer at surgery, which was common to all ABLs (FEA, ADH, ALH or LCIS). NOMAT is based on three preoperative variables: age, complete radiological target excision after biopsy (yes/no) and focus size (≤ or > 15mm) (10). All these criteria have indeed constantly been reported as risk factors for breast cancer in numerous studies (11). NOMAT's performance was promising in the development cohort, with an area under the receiving operating characteristics (ROC) curve (AUC) of 0.72 (95 % confidence interval [CI], 0.63–0.82). For an optimal threshold of 20% of predicted probability of cancer using NOMAT, the sensitivity (Se), specificity (Sp), positive and negative predictive values (PPV and NPV) were 78%, 66%, 36%, and 90%, respectively. Our model was also externally validated by another academic team on a series of 151 patients, with an AUC of 0.65 (0.63-0.82) (14).

The main objective of the present study was to evaluate NOMAT's performance among women diagnosed with ABL at breast biopsy in a multicenter prospective setting, and we explored the performance in the population restricted to patients with ADH

Patients and Methods

Patients

Women who had a recent diagnosis of ABL on image-guided biopsy for breast microcalcifications were included in this trial between July 7, 2015 and May 29, 2018, in 18 French centers. A French Ethics Committee (Comité de Protection des Personnes Ile de France) approved this study on March 11, 2015 (2015-A00045-44). The trial was registered at clinicaltrials.gov under the number NCT: NCT02523612. Women aged 18 or older were eligible for the study if diagnosed with an ABL (including flat epithelial atypia [FEA], atypical ductal [ADH] or lobular hyperplasia [ALH] and lobular carcinoma in situ [LCIS] of classic type) at percutaneous breast biopsy performed for mammographic microcalcifications. All women gave a written informed consent before study participation. Exclusion criteria were the presence of a mass lesion on ultrasound imaging; the presence of DCIS or invasive carcinoma on the biopsy specimen; a past personal history of breast cancer; a previous radiation therapy of the breast; a known germline predisposing mutation. Pregnant or breastfeeding women, patients with limitation of freedom or under guardianship, and patients unable to comply with the necessary follow-up for geographical, social or psychological reasons were also excluded. Multifocal lesions were included, and the lesion at higher risk was considered as the main lesion for characterization.

Baseline epidemiological data were collected. Locally assessed imaging features were collected as well: American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) classification of the baseline mammographic image, visual mammographic breast density (BI-RADS), size of the target lesion before biopsy (mm), size of the biopsy needle, number of cores retrieved and completeness of the microcalcifications' removal during biopsy.

All women included in the study underwent surgical image-guided excision of the microcalcifications (lumpectomy) by their local surgeon.

Pathology

Diagnostic biopsy:

Initial diagnosis and inclusion were based on local pathological analysis to reflect real-practice conditions. There was no central review prior to inclusion.

Sections were obtained at a minimum of three levels for 11G needle biopsies and five levels for 8G needle biopsies. All slides were stained with hematoxylin–eosin-safran (HES) and analyzed by local breast pathologists according to routine practice. Epithelial atypical lesions were classified according to the World Health Organization 2012 classification as FEA, ADH, ALH, or LCIS of classic type. When several lesions were associated, the diagnosis was based on the predominant lesion but associated lesions were also documented. The number of foci with atypia and the size of the largest foci were specified for FEA and ADH. For ALH and LCIS, the percentage of lobular and ductal lesions with atypia was reported.

Surgical specimen:

On the surgical specimen, the presence of post-biopsy scar, of any residual atypical lesions, and of any malignancy either close to or at a distance from the scar was recorded. If a cancer was observed, the histological type (invasive ductal of no special type, lobular

carcinoma, invasive special type, DCIS, pleomorphic LCIS), histological grade, tumor size and margin status were specified. For invasive cancers, estrogen and progesterone receptors (ER and PR), and human epidermal growth factor receptor 2 (HER2) status were also assessed. New surgeries and subsequent treatment were reported.

The patients were classified as bearing "cancer" on the final surgical sample if either an invasive adenocarcinoma or a DCIS or pleomorphic LCIS was retrieved, whatever the extent.

Individual assessment of NOMAT model

The following variables are included in the NOMAT model: age (years), complete radiological target excision after biopsy (yes/no) and initial radiological microcalcifications focus size (≤ or > 15mm).

For each woman, the probability of cancer at surgery was predicted using the NOMAT multivariable logistic model whose parameters were slightly optimized from a previous publication (10) and locked before the start of the current NOMAT01 study : logit(P[breast cancer = 1 | age, focus size > 15, excision target] = -4.42 + 0.05*Age + 1.01*focus size > 15 -0.87*excision target).

Statistical analysis

To evaluate the performance of the NOMAT model, a sample size of 300 patients was deemed necessary. The null hypothesis was given by H0: NPV (=true negatives/(true negatives + false negatives)) = 0.80 vs. the alternative hypothesis HA: NPV > 0.80. We assumed that 50% of the patients would be classified in the low-risk group (i.e. below the pre-specified NOMAT cut-off of 20% on the scale of the predicted probability of cancer). The inclusion of 300 women would provide 84% power to reject the null hypothesis at a one-sided significance level of 5% using an empirical estimate of variance when the true NPV is equal to 0.88 under the alternative hypothesis.

To assess the performance of the NOMAT model, NPV, Se (=true positives/(true positives + false negatives)), Sp (=true negatives/(true negatives + false positives)), PPV (=true positives/(true positives + false positives)) were calculated for the 20% cutoff on the scale of predicted probability of cancer and were provided with 95% CIs. The discrimination of the model was assessed by an estimation of the AUC, and provided with a 95% CI. Additionally, because of its particular relevance given the recent changes in practices in women with non-ADH lesions, an unplanned subgroup analysis for patients with ADH lesions was performed. The calibration of the model was assessed using a calibration plot and complemented with a calibration intercept and slope (15). Statistical analyses were performed on SAS version 9.4 and R version 3.5.0 (using ROCR and gbm packages).

Results

Patient characteristics

Three hundred patients were included in 18 centers. Eleven patients were excluded from the analysis (Figure 1). Out of the 289 remaining patients, 2 patients had missing data for lesion size on mammography, leaving 287 patients evaluable for the assessment of the prediction model's performance.

Patients' main characteristics are reported in Table 1. Median age was 52 years (range: 32-81). ABLs at biopsy were distributed as follows: 97 ADH alone (34%), 16 ALH alone (6%), 7 classic-type LCIS (2%), 49 FEA (17%), 113 mixed lesions (39%), and 5 with undefined atypical lesions (2%). ADH was present in 195/287 patients (97 alone and 98 mixed).

Surgical findings and cancer characteristics

At the time of surgery, 174 patients (61%) had residual ABL, while 113 patients (39%) did not. The residual ABL was FEA in 99 (57%), ADH in 95 (55%), ALH in 44 (25%) and classic-type LCIS in 14 patients (8%). Among the 287 patients operated, 55 cancers were found at surgery (19%) including 41 DCIS, 2 pleiomorphic LCIS and 12 invasive carcinoma (10 of no special type, 1 tubular, and 1 lobular; all grade 1). All 12 invasive cancers were node-negative-11 stage 1 and one stage 2, with a median tumor size of 7 mm [range: 1; 25]. Ten out of 11 invasive carcinoma were ER+/PR+/HER2-, and 1 carcinoma was too small for immunohistochemistry characterization. The margins were free of disease for invasive carcinoma in 11 out of 12 patients. For DCIS or pleiomorphic LCIS, the median tumor size was 17 mm [range: 3-70]. The 43 DCIS/pleiomorphic LCIS were distributed as follows: 16 low, 23 intermediate and 3 high-grade lesions (one missing grade). For DCIS (41 patients), necrosis was observed for 8 patients (8/41, 20%) and median size of lesion was 15.5mm (Q1-Q3: 7.5–30.0 / range: 3–70).

Thirty-three patients (12%) had a secondary surgery (18 simple lumpectomy, 3 lumpectomy + sentinel lymph node dissection [SND], 1 lumpectomy + axillary lymph node dissection [ALND], 10 mastectomy + SND and 1 mastectomy + ALND). Forty-one patients (14% of the whole cohort, 75% of cancer patients) received a complementary treatment (39 radiotherapy, 1 chemotherapy, and 8 endocrine therapy).

The upgrading on surgical excision according to pathological subtype on biopsy is reported in Table 2. This rate was 25% (49/195) for patients with ADH alone or mixed with other ABL. No women with isolated FEA (n=49) at the biopsy exhibited cancer on surgical excision.

NOMAT model's performance

The AUC of the NOMAT model in the whole ABL cohort was 0.64 (95% CI, 0.58 to 0.69) (Figure 2). The AUC of the NOMAT model for patients with ADH (alone or associated, n=195) was 0.63 (95% CI, 0.56 to 0.70).

Regarding the primary outcome, using the pre-specified threshold of 20% on the predicted probability of cancer among the whole cohort (n=287), 39 patients were predicted to have cancer, with a corresponding Sensibility, Specificity, Positive Predictive Value and Negative Predictive Value of 20% (95% CI, 9% to 30%), 88% (95% CI, 84% to 92%), 28% (95% CI, 14% to 42%) and 82% (95% CI, 77% to 87%), respectively.

When applying an exploratory threshold of 10% on the predicted probability of cancer, 120 patients were predicted to have cancer (Figure 3), with a corresponding Se, Sp, PPV and NPV of 65% (95% CI, 52% to 78%), 67% (95% CI, 61% to 73%), 31% (95% CI, 23% to 39%) and 89% (95% CI, 84% to 94%), respectively. By using the model with the 10% risk threshold, 58% of the patients could have avoided surgery with only 2 missed invasive cancers but 16 missed DCIS or pleiomorphic LCIS (Figure 3A).

When examining the nuclear grade of DCIS or pleiomorphic LCIS according to this 10% cut-off of predicted probability of cancer, no high-grade would have been missed and 61% (14/23) of intermediate-grade DCIS would have been identified as high-risk.

For the ADH population (ADH alone or associated with other ABL) (n=195), when applying the pre-specified threshold of 20% on the predicted probability of cancer, 29 patients were predicted to have cancer, with a corresponding Sensibility, Specificity, Positive Predictive Value and Negative Predictive Value of 22% (95% CI, 11% to 34%), 88% (95% CI, 82% to 93%), 38% (95% CI, 20% to 56%) and 77% (95% CI, 71% to 83%), respectively. When applying an exploratory threshold of 10% on the predicted probability of cancer among this ADH population, 90 patients were predicted to have cancer, with a corresponding Sensibility, Specificity, Positive Predictive Value and Negative Predictive Value of 67% (95% CI, 54% to 80%), 61% (95% CI, 53% to 69%), 37% (95% CI, 27% to 47%) and 85% (95% CI, 78% to 92%), respectively. By using the model with the exploratory 10% risk threshold, 105/195 (54%) patients could have avoided surgery with only 2 missed invasive cancers but 14 missed DCIS or pleiomorphic LCIS (Figure 3B).

Regarding the calibration of the NOMAT model, small risks of cancer were moderately overestimated, whereas large risks of cancer were severely underestimated (calibration slope equal to 0.49 [95% CI, 0.19 to 0.80], with an intercept equal to 0.65 [95% CI, 0.33 to 0.96]) (Figure 4).

Discussion

This study is the first prospective multicentric validation of a model to predict the probability of cancer for women with ABL on biopsy. The NOMAT model performed modestly well in this multicentric prospective validation cohort. The trial reached its primary objective: at a pre-specified threshold of 20% predicted probability of cancer, the NPV of NOMAT was beyond 80%, precisely 82% (95% CI, 77% to 87%). The model's discrimination in this validation series appeared lower compared to the initial study results (AUC 0.64 vs. 0.72), which is often observed in validation studies. The performance of the model similar in the analysis restricted to patients with ADH (alone or associated, AUC equal to 0.63). When applying an exploratory cut-off of 10%, the NPV was estimated to be 89% (95% CI, 84% to 94%), which appeared more relevant. Indeed, this threshold would lead to avoid 58% of surgeries (and 54% in patients with ADH) with only 2 missed invasive cancers and 16 missed DCIS or pleiomorphic LCIS (none being high-grade).

Several international studies currently question the management of low-grade DCIS and propose an alternative of follow-up for those lesions (LORD NCT02492607, LORIS ISRCTN 27544579 and COMET (6-8)). In this respect, a model allowing the avoidance of surgery for ABL should primarily identify the risk for invasive cancer and high-grade DCIS. This is what we obtain with NOMAT, based on current results, which appears as a useful decisional tool for clinicians face to ABL on biopsy.

International recommendations have been published regarding each type of benign breast lesion or B3 lesions. Minimally invasive management of B3 lesions (except ADH) with vacuum-assisted biopsy (VAB) was considered an appropriate alternative to first-line open surgical excision in most cases, but with more frequent follow-up, especially for ALH (9). If decision for surgery is based essentially on pathology analysis, a key question is how to be sure of the type of B3 lesion. In the literature review by Mooney et al. (11), 23% of ADH cases [0-62%], 9% of ALH cases [0-67%], and 18% of LCIS cases [0-60%] were upgraded on average at surgery (16-20). Indeed, the trend over the last several years is to not excise ALH and classic LCIS identified on core biopsy when there are concordant imaging findings. This recommendation is now in the NCCN and international guidelines (9, 21). In those guidelines, the major unresolved and clinically relevant question is which women with ADH on core biopsy can be spared excision.

However, the pathological reproducibility is lower for ABL as compared to cancer. Elmore et al. studied diagnostic concordance among pathologists interpreting breast biopsy specimens (22): for invasive carcinoma cases, 96% (95% CI, 94%-97%) were concordant, and 4% (95% CI, 3%-6%) underinterpreted; among DCIS cases, 84% (95% CI, 82%-86%) were concordant, 3% (95% CI, 2%-4%) overinterpreted, and 13% (95% CI, 12%-15%) underinterpreted; while among ABL cases, 48% (95% CI, 44%-52%) were concordant, 17% (95% CI, 15%-21%) overinterpreted, and 35% (95% CI, 31%-39%) underinterpreted. Moreover, lesions are often mixed (in our study, nearly 40% of patients). These elements may explain why it is difficult to rely solely on the pathological report of the biopsy to decide between surgery or follow-up. Our model does not integrate B3 subtyping, therefore saving a mandatory central review of the lesions.

While several prediction models aiming to avoid surgery have been proposed previously, most concerned specific subtypes and none has been prospectively validated so far. Several models have included the number of ADH foci (4, 11, 23-25). However, in our multicentric study, although we collected this factor prospectively, it was missing in 48 out of 195 lesions with ADH (alone or mixed with other lesions), illustrating the limitation for daily

use without expert analysis. Two machine learning models have been developed recently. A random forest model was developed within a single-institution series of 1006 patients, with as most important predictors, age and histology (26). This model has not been validated yet. Harrington et al. developed and evaluated six machine learning models on a small patient series of 124 women. They reported that patient's age at time of the biopsy together with lesion size were the most important predictors of upgrade, as in our model (27). Our model can be used as a tool to inform women regarding their probability of cancer within a biopsied microcalcification spot and help decision-making between surgical or follow-up strategies. Such decisions undoubtedly need to be adapted to each patient's anxiety level, tolerance to uncertainty and adherence to follow-up. Patients' preference with adapted evaluation of risks shall guide the final decisions.

Some weaknesses can be acknowledged for this study. All patients were operated and this study did not intend to evaluate the long-term risk of cancer in the presence or absence of surgery. It provides the scientific community with a validated, short-term, surgical-decision model. No central pathological review of the slides was implemented, but our objective was to develop a daily practice tool, and expert review is not always available for these frequent ABLs. Very few patients had an MRI whereas, for some authors, this imaging could add valuable information in the investigation of breast lesions (28, 29). With an exploratory threshold of 10% in our validation study, surgery would have been avoided in 58% of women with only 2 missed invasive cancer and 16 missed DCIS or pleiomorphic LCIS (but not high-grade). Thus, the use of NOMAT safely allows avoidance of surgery at a preferred threshold of 10% probability. This threshold needs to be confirmed in future studies. Prospective follow-up data of women who were spared surgery shall be gathered to confirm the long-term safety of such an approach. Finally, the model's discrimination and calibration could be improved. Indeed, a secondary aim of the NOMAT01 study is to use biomarkers to further improve the performances of the model. Imperfect calibration may be linked to the differences of case mix between the development cohort and the present one. Regarding the three predictors of the NOMAT01 model the current study included slightly younger patients; of note however is that the distribution of ABL subtypes was rather different.

Also some patients had multiple lesions, and the decision to pursue the "highest risk" lesion can be arguable. Most of the time this "highest risk" lesion was ADH and, in the literature, to decide on patient management, the existence of ADH usually determines the choice for surgery (9, 21).

In conclusion, we conducted the first multicentric prospective validation of a model that predicts the probability of breast cancer at surgery among women with ABL. The model performed modestly well in this multicentric prospective validation cohort and may need some recalibration. The use of a model such as NOMAT allows the safe avoidance of surgery at the exploratory threshold of 10% probability of cancer. The NOMAT model can help inform women and guide decisions between surgical excision or follow-up and can be applied to all ABLs identified on microcalcifications.

Acknowledgements

We thank all women who kindly accepted to participate in this study. We thank the research teams of all participating centers, and Nathalie Bouvet for data management. We thank the French National Cancer Institute for funding this trial.

Funding

This trial was supported by a French National Grant after expert reviews (PHRC National Cancer 2014 Reference INCA RECF2719)

References

1. Berry DA, Cronin KA, Plevritis SK, Fryback DGet al: Cancer Intervention and Surveillance Modeling Network (CISNET) Collaborators. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 353:1784-92, 2005

2. Plevritis SK, Munoz D, Kurian AW, et al: Association of Screening and Treatment With Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012. JAMA. 319:154-164, 2018

3. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness. N Engl J Med. 2016; 375:1438-1447.

4. Hartmann LC, Sellers TA, Frost MH et al. Benign breast disease and the risk of breast cancer. NEngl J Med 353:229–37, 2005

5. Mastropasqua MG, Viale G. Clinical and pathological assessment of high-risk ductal and lobular breast lesions: What surgeons must know. Eur J Surg Oncol. 43:278-284, 2017

6. Elshof LE, Tryfonidis K, Slaets L, et al: Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ - The LORD study. Eur J Cancer. 51:1497-510, 2015

7. Francis A, Thomas J, Fallowfield L, et al: Addressing overtreatment of screen detected DCIS; the LORIS trial. Eur J Cancer. 51:2296-303, 2015

8. Hwang ES, Hyslop T, Lynch T, et al: The COMET (Comparison of Operative versus Monitoring and Endocrine Therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). BMJ Open. 9:e026797, 2019

9. 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.;174:279-296, 2019

10. Uzan C, Mazouni C, Ferchiou M, et al: A Model to Predict the Risk of Upgrade to Malignancy at Surgery in Atypical Breast Lesions Discovered on Percutaneous Biopsy Specimens. Ann Surg Oncol. 20:2850-7, 2013

11. Mooney KL, Bassett LW, Apple SK. Upgrade rates of high-risk breast lesions diagnosed on core needle biopsy: a single-institution experience and literature review. Mod Pathol. 29:1471-1484, 2016

12. Rageth CJ, Rubenov R, Bronz C, et al: Atypical ductal hyperplasia and the risk of underestimation: tissue sampling method, multifocality, and associated calcification significantly influence the diagnostic upgrade rate based on subsequent surgical specimens. Breast Cancer. 26:452-458, 2019

13. Bendifallah S, Defert S, Chabbert-Buffet N, et al: Scoring to predict the possibility of upgrades to malignancy in atypical ductal hyperplasia diagnosed by an 11-gauge

vacuum-assisted biopsy device: an external validation study. Eur J Cancer. 48:30-6, 2012

14. Linsk A, Mehta TS, Dialani V, et al: Surgical upgrade rate of breast atypia to malignancy: An academic center's experience and validation of a predictive model. Breast J. 24:115-119, 2018

15. Van Calster B, Nieboer D, Vergouwe Y, et al: A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol. 74:167-76, 2016

16. McGhan LJ, Pockaj BA, Wasif N, et al: Atypical ductal hyperplasia on core biopsy: an automatic trigger for excisional biopsy? Ann Surg Oncol. 19:3264-3269, 2012

17. Ko E, Han W, Lee JW, et al. Scoring system for predicting malignancy in patients diagnosed with atypical ductal hyperplasia at ultrasound-guided core needle biopsy. Breast Cancer Res Treat. 112:189-195, 2008

18. Buckley E, Sullivan T, Farshid G, et al: Risk profile of breast cancer following atypical hyperplasia detected through organized screening. Breast. 24:208-212, 2015

19. Mazzola E, Coopey SB, Griffin M, et al. Reassessing risk models for atypical hyperplasia: age may not matter. Breast Cancer Res Treat. 165:285-291, 2017

20. Vierkant RA, Degnim AC, Radisky DC, et al. Mammographic breast density and risk of breast cancer in women with atypical hyperplasia: an observational cohort study from the Mayo Clinic Benign Breast Disease (BBD) cohort. BMC Cancer. 17:84, 2017

21. Bevers TB, Helvie M, Bonaccio E, et al. Breast Cancer screening and diagnosis, version 3.2018, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw. 2018;16(11):1362–89.

22. Elmore JG, Longton GM, Carney PAet al: Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA. 313:1122-32, 2015

23. Forgeard C, Benchaib M, Guerin N, et al: Is surgical biopsy mandatory in case of atypical ductal hyperplasia on 11-gauge core needle biopsy? A retrospective study of 300 patients. Am J Surg. 196:339-45, 2008

24. Khoury T, Li Z, Sanati S, et al: The risk of upgrade for atypical ductal hyperplasia detected on magnetic resonance imaging-guided biopsy: A study of 100 cases from four academic institutions. Histopathology 68:713-721, 2016

25. Peña A, Shah SS, Fazzio RT, et al: Multivariate model to identify women at low risk of cancer upgrade after a core needle biopsy diagnosis of atypical ductal hyperplasia. Breast Cancer Res Treat 164:295-304, 2017

26. Bahl M, Barzilay R, Yedidia AB, et al: High-risk breast lesions: A machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision. Radiology 286:810-818, 2018

27. Harrington L, diFlorio-Alexander R, Trinh K, MacKenzie T, Suriawinata A, Hassanpour S. Prediction of Atypical Ductal Hyperplasia Upgrades Through a Machine Learning

Approach to Reduce Unnecessary Surgical Excisions. JCO Clin Cancer Inform. 2:1-11, 2018

28. Speer ME, Huang ML, Dogan BE, Adrada BE, Candelaria RP, Hess KR, Hansakul P,Yang WT, Rauch GM. High risk breast lesions identified on MRI-guided vacuum-assisted needle biopsy: outcome of surgical excision and imaging follow-up. Br J Radiol. 2018 Oct;91(1090):20180300.

29. Tsuchiya K, Mori N, Schacht DV, Sheth D, Karczmar GS, Newstead GM, Abe H. Value of breast MRI for patients with a biopsy showing atypical ductal hyperplasia (ADH). J Magn Reson Imaging. 2017 Dec;46(6):1738-1747.

List of figures and tables

Table 1: Description of patients' characteristics at inclusion (n=287)

Table 2: Upgrading on surgical excision according to pathological subtype on biopsy (n=287)

Figure 1: Flowchart of the NOMAT01 study

Figure 2: ROC curve of the NOMAT model (n=287)

Figure 3: Predicted probability of breast cancer (n=287)

Figure 4: Calibration plot of the NOMAT model (n=287)

Table 1: Description of patients' characteristics at inclusion (n=287)

Age at inclusion	N	287	
	Median (Min ; Max)	52 (32 ; 81)	
Family history of breast cancer (first and second degree)	N	284	
	No	166	58.5%
	Yes	118	41.5%
Personal history of breast biopsy	N	286	
	No	238	83.2%
	Yes	48	16.8%
Personal history of atypical lesion	N	282	
	No	266	94.3%
	Yes	16	5.7%
	Ipsilateral	10	62.5%
	Contralateral	6	37.5%
Menopausal	N	283	
	No	139	49.1%
	Yes	144	50.9%
Hormone Replacement Therapy	N	144	
	No	103	71.5%
	Yes	41	28.5%

Table 2: Upgrading on surgical excision according to pathological subtype on biopsy (n=287)

All patients	Cancer on surgical excision (= DCIS, pleiomorphic LCIS, IC) (n=55) n (%)	No cancer on surgical excision (n=232) n (%)
ADH alone (n=97)	27 (27.8)	70 (72.2)
ALH alone (n=16)	1 (6.25)	15 (93.75)
LCIS alone (n=7)	1 (14.3)	6 (85.7)
FEA (n=49)	0 (0.0)	49 (100.0)
Mixed lesions (n=113)	23 (20.4)	90 (79.6)
Undefined atypical lesions (n=5)	3 (60.0)	2 (40.0)
Subgroup		
ADH alone and mixed (n=195)	49 (25.1)	146 (74.9)

Abbreviations: ADH=atypical ductal hyperplasia – ALH=atypical lobular hyperplasia – LCIS=lobular carcinoma in situ – FEA=flat epithelial atypia

Figure 1: Flowchart of the NOMAT01 study

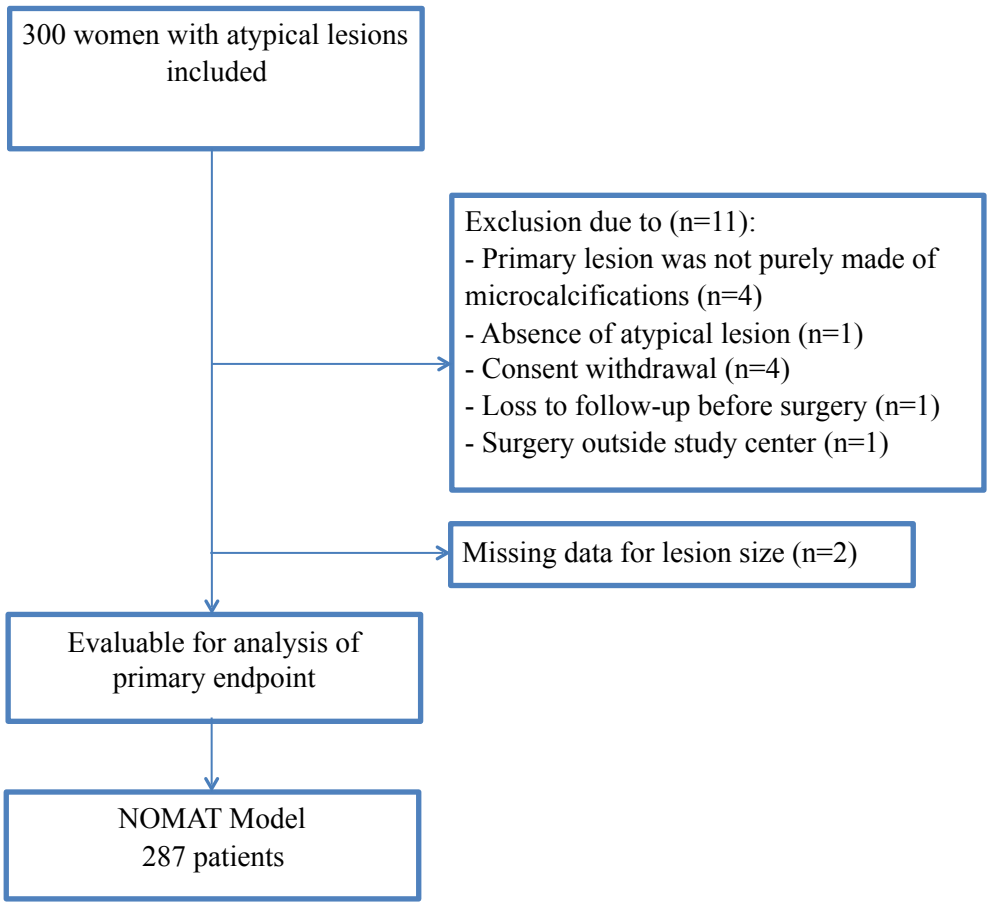


Figure 2: ROC curve of the NOMAT model

The Area Under the Receiving Operating Characteristics Curve (AUC) of the NOMAT model in the whole cohort (n=287) was 0.64 (95% CI, 0.58 to 0.69).

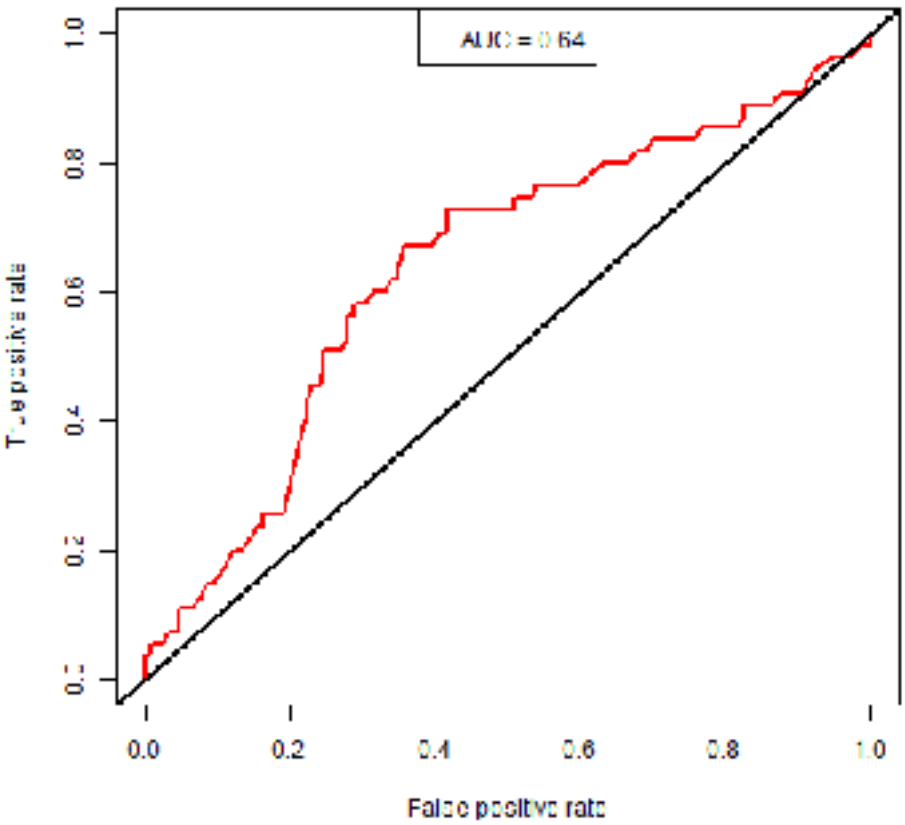


Figure 3: Predicted probability of breast cancer for the whole cohort (n=287, Figure 3A) and for patients with ADH (alone or associated, n=195, Figure 3B)

The patients were classified as bearing "cancer" on the final surgical sample if either an invasive adenocarcinoma or a DCIS or pleomorphic LCIS was retrieved, whatever the extent.

Figure 3A

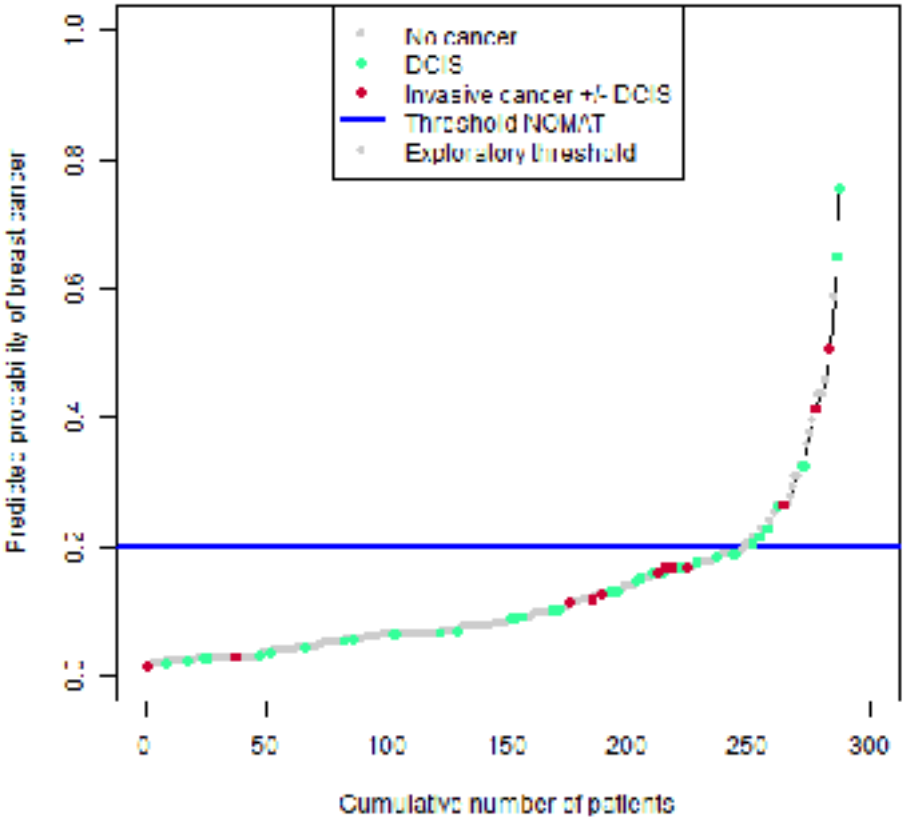


Figure 3B

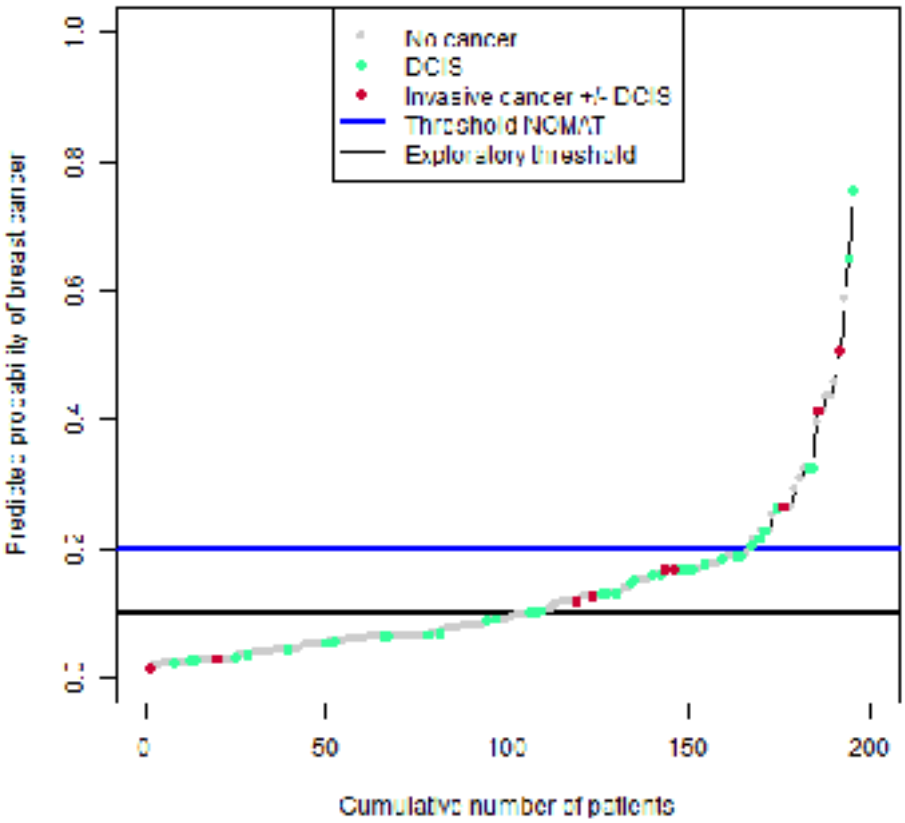
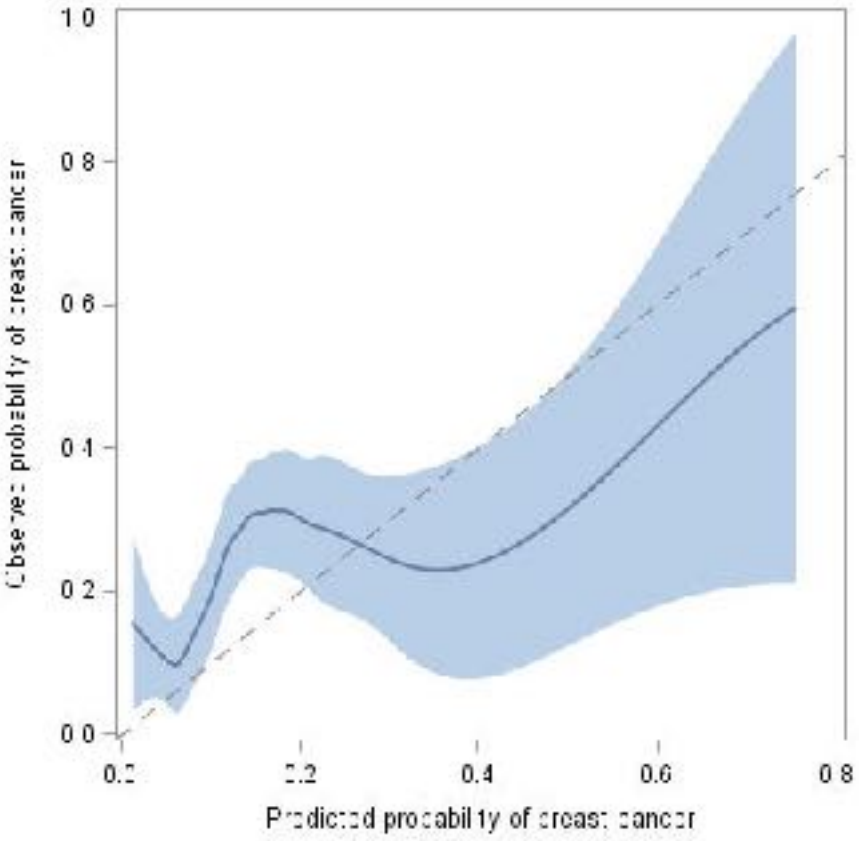

No cancer
DCIS
Invasive cancer +/- DCIS
Threshold NOMAT
Exploratory threshold
Predicted probability of breast cancer
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
Cumulative number of patients

Figure 4: Calibration plot of the NOMAT model (n=287)

Blue line corresponds to a LOESS fit (with 95% confidence bounds), the dotted line to the theoretical perfect calibration.