EUROPEAN JOURNAL OF CANCER 48 (2012) 30–36

Available at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.ejconline.com

Current Perspective

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

S. Bendifallah [a,b,*], S. Defert [c], N. Chabbert-Buffet [a], N. Maurin [d], J. Chopier [d], M. Antoine [e], C. Bezu [a], D. Touche [f], S. Uzan [a,e], O. Graesslin [c], R. Rouzier [a,b,g]

[a] Department of Obstetrics and Gynecology, Tenon APHP University Hospital, 75020 Paris, France
[b] ER2, Pierre and Marie Curie University, Paris, France
[c] Department of Obstetrics and Gynaecology, Institute Alix de Champagne University Hospital, 51092 Reims, France
[d] Department of Pathology, Tenon APHP University Hospital, Paris, France
[e] Department of Radiology, Tenon APHP University Hospital, Paris, France
[f] Breast Unit, Institute Jean Godinot, Reims, France
[g] INSERM-UMR_S 938, Pierre and Marie Curie University, Paris, France

ARTICLE INFO

Article history:
Available online 17 November 2011

Keywords:
Atypical ductal hyperplasia
Mathematical models
Underestimation rate
Upgrading rate
Scoring system
Vacuum-assisted biopsy

ABSTRACT

Background: Ko's scoring system was developed to predict malignancy upgrades in patients diagnosed with atypical ductal hyperplasia by core needle biopsy. The Ko algorithm was able to identify a subset of patients who were eligible for exclusively clinical follow-up. The current study statistically investigated the patient outcomes to determine whether this scoring system could be translated and used safely in clinical practice.

Methods: We tested the statistical performance of the Ko scoring system against an external independent multicentre population. One hundred and seven cases of atypical ductal hyperplasia diagnosed by an 11-gauge biopsy needle were available for inclusion in this study. The discrimination, calibration and clinical utility of the scoring system were quantified. In addition, we tested the underestimation rate, sensitivity, specificity, and positive and negative predictive values according to the score threshold.

Results: The overall underestimation rate was 19% (20/107). The area under the receiver operating characteristic curve for the logistic regression model was 0.51 (95% confidence interval: 0.47–0.53). The model was not well calibrated. The lowest predicted underestimation rate was 11%. The sensitivity, specificity, positive predictive value, and negative predictive values were 90%, 22%, 20%, and 89%, respectively, according to the most accurate threshold proposed in the original study.

Conclusion: The scoring system was not sufficiently accurate to safely define a subset of patients who would be eligible for follow-up only and no additional treatment. These results demonstrate a lack of reproducibility in an external population. A multidisciplinary approach that correlates clinicopathological and mammographic features should be recommended for the management of these patients.

* *Corresponding author:* Tel.: +33 (0)1 56 01 68 76/68 49; fax: +33 (0)1 56 01 60 62.
E-mail address: sofiane.bendifallah@yahoo.fr (S. Bendifallah).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.08.011

1. Introduction

Population-based mammography screening has resulted in increased detection of suspicious, non-palpable lesions that require further histopathological assessment. Ultrasound-guided needle biopsy (14–16 gauge) or vacuum-assisted (11–14 gauge) breast biopsy (VABB) systems have become widely-used alternatives to open surgical biopsy.[1–3] Atypical ductal hyperplasia (ADH) of the breast, which is discovered in 2–11%[4,5] of cases, is histopathologically defined as either (i) a hyperplastic lesion with some cytological features of low-grade ductal carcinoma in situ (DCIS) that lacks the overall characteristic architectural growth pattern of DCIS, or (ii) a lesion with the classic cytological and architectural features of low-grade DCIS that is confined to ducts and measures less than 2 mm.[6] The difficulty in achieving acceptable levels of concordance between pathology results from image-guided biopsy (IGB) and surgical excision is a major practical concern.[7,8] Due to the risk of underestimating or upgrading the diagnosis (meaning that DCIS or invasive cancer are present), surgical excision is an accepted option for all women diagnosed with ADH. Various strategies have been unsuccessfully developed to improve cancer detection and the risk of underestimation, including revising the definition criteria, changing the device size, and testing clinical, radiological or pathological factors.[9–13] A risk of upgrade of 2% or less has been suggested by the American College of Radiology[14] to be safe for proposing exclusive follow-up breast imaging in certain cases. Based on a population of ADH cases diagnosed by ultrasound-guided core needle biopsy (CNB), in 2007, Ko et al.[15] developed a logistic regression model as an algorithm for scoring the possibility of predicting malignancy upgrade using a combination of five independent factors. The accuracy of the model was tested. The area under (AUC) the receiver operating characteristic (ROC) curve was 0.90 (95% confidence interval, 0.83–0.97) and 0.85 (95% CI, 0.74–0.95) in the study (74 patients) and validation datasets (54 patients), respectively. Because the study indicated a reported sensitivity and negative predictive value of 100% and no cases were upgraded amongst those with a score of 3.5 or less, the authors concluded that this subset of patients should be eligible for non-invasive management.[15] These relevant findings support the hypothesis that the scoring system should be applicable to another population. To our knowledge, no external validation of this tool has been published. Therefore, in the current report, we evaluated the performance and clinical utility of Ko's scoring system[15] in our population of samples from 11-gauge VABB to determine whether this system could be used in clinical practice.

2. Materials and methods

2.1. Data selection

A multicentre search of the medical databases at Tenon APHP University Hospital and the Institute Alix de Champagne University Hospital to identify only ADH cases diagnosed by imaging-guided biopsies (11-gauge vacuum-assisted biopsy device) and followed by surgical excision between January 2003 and December 2010 revealed 229 cases. Amongst these, 13 cases in which the ADH was associated with malignant lesions (i.e. invasive carcinoma or DCIS) upon biopsy and 109 cases in which the absence of one relevant (radiographic, pathological or clinical) criterion prevented the use of the Ko nomogram for scoring were excluded. The details regarding the missing parameters are reported in Fig. 1. Therefore, 107 (46.7%) cases were eligible for the current validation study. Demographical data, imaging, biopsy and open surgical pathology results were collected for each patient (Table 1).

2.2. Mammography and biopsy evaluation

Biopsy procedures were performed for mammographically detected microcalcifications ($n = 100$), architectural distortions and/or suspicious asymmetric densities ($n = 6$) and palpable mass lesions ($n = 1$) by breast intervention radiologists using an 11-gauge vacuum-assisted biopsy device (Mammotome; Ethicon Endo-Surgery, Cincinnati, Ohio). Each pre-biopsy mammogram was independently reviewed to categorise the lesions according to the mammographic BI-RADS category[14] (categories 3–5) and to measure the maximum mammographic lesion diameter. The percutaneous biopsy specimens and surgery slides of the 107 cases were collected from the electronic medical records and specifically re-reviewed for this analysis. The histological slides were interpreted at either Tenon's University Hospital or at the Institute Alix de Champagne University Hospital by experienced pathologists, and were diagnosed according to the diagnostic criteria of the revised 2003 World Health Organization guidelines and classification.[5] Lesions yielding ADH at biopsy and DCIS or carcinoma at surgery were recorded as ADH underestimations.

All excisions were guided by preoperative wire-localisation, and pathological analysis was performed on the retrieved vacuum cavity. The absence of malignancy was considered to be coincidental and therefore predictable by the Ko model. This consideration does not warrant the continuum between ADH and malignancy but is important to test the accuracy of the Ko model.

2.3. Development of the Ko scoring system

The five independent predictors of malignancy included in the scoring system were selected based on P values <0.05 in the multivariate analyses. The multiple of 0.5 nearest to the β coefficient obtained for each significant factor from the multivariate logistic regression model was assigned for each factor to build the algorithm. A score of 2.0 was assigned for a palpable lesion and microcalcification on mammography; 3.5 was assigned for calcifications >1.5 cm, focal ADH ($\leqslant$1 duct and $\leqslant$1 mm), and age >50 years. The scores for each significant factor were then added, resulting in a total score for each patient. The final scores ranged from 0 to 14.5. A score $\leqslant$3.5 designated a subset of patients with ADH lesions that were defined as 'probably benign' and could be safely followed non-operatively rather than surgically excised.[15]