ORIGINAL ARTICLE

Advancement study of CancerMath model as prognostic tools for predicting Sentinel lymph node metastasis in clinically negative T1 breast cancer patients

Raffaella Massafra¹, Domenico Pomarico¹, Annarita Fanizzi¹, Francesco Campobasso², Vittorio Didonna¹, Agnese Latorre³, Annalisa Nardone⁴, Irene Maria Pastena⁵, Pasquale Tamborra¹, Vito Lorusso², Daniele La Forgia⁶

¹Department of Health Physics, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Bari, Italy. ²Department of Economics and Finance, University of Bari, Italy. ³Medical Oncology Unit, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Bari, Italy. ⁴Radiotherapy Unit, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Bari, Italy. ⁵Pathological Anatomy Unit, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Bari, Italy. ⁶Department of Breast Radiology, I.R.C.C.S. Istituto Tumori “Giovanni Paolo II”, Bari, Italy.

Summary

Purpose: Sentinel lymph node biopsy (SLNB) is an invasive surgical procedure and although it has fewer complications and is less severe than axillary lymph node dissection, it is not a risk-free procedure. Large prospective trials have documented SLNB that it is considered non-therapeutic in early stage breast cancer.

Methods: Web-calculator CancerMath (CM) allows you to estimate the probability of having positive lymph nodes valued on the basis of tumour size, age, histologic type, grading, expression of estrogen receptor, progesterone receptor. We collected 595 patients referred to our Institute resulting clinically negative T1 breast cancer characterized by sentinel lymph node status, prognostic factors defined by CM and also HER2 and Ki-67. We have compared classification performances obtained by online CM application with those obtained after training its algorithm on our database.

Results: By training CM model on our dataset and using the same feature, adding HER2 or ki67 we reached a sensitivity median value of 71.4%, 73%, 70.4%, respectively, whereas the online one was equal to 61%, without losing specificity. The introduction of the prognostic factors Her2 and Ki67 could help improving performances on the classification of particularly type of patients.

Conclusions: Although the training of the model on the sample of T1 patients has brought a significant improvement in performance, the general performance does not yet allow a clinical application of the algorithm. However, the experimental results encourage future developments aimed at introducing features of a different nature in the CM model.

Key words: sentinel lymph node, T1 breast cancer, clinical decision support system, clinically negative lymph node, CancerMath

Introduction

T1 stage breast cancer has become the most frequently diagnosed invasive form of this disease. In fact, with the introduction of breast cancer programs, the number of patients diagnosed with early-stage cancer and reduced axillary lymph node involvement has increased [1,2].

In existing clinical practice, in the absence of lymph node abnormalities detectable by clinical examination or imaging [3], the procedures recommend the removal of the first axillary draining lymph nodes to investigate the positivity of some lymph nodes identified by injection of a radionu-
These lymph nodes called “sentinel” and recognized with a procedure known as sentinel node and occult lesion localization (SNOLL), are objects of a second intra-operative pathological examination, whose experimental technique (one-step nucleic acid amplification; OSNA) goes above clinical and imaging limitations, reaching a sufficient performance required for clinical application, at the price of an expensive and time-consuming procedure, currently with sensitivity between 87.5-100% and specificity between 90.5 and 100% [5-8].

Although the OSNA technique is time-consuming and expensive, it is still the intraoperative exam with the currently best results. Nevertheless, sentinel lymph node biopsy (SLNB) is an invasive surgical procedure and although it has fewer complications and is less severe than axillary lymph node dissection [9,10], it is not a risk-free procedure.

Large prospective trials have documented SLNB complications including allergic reactions, wound infection, seroma, paresthesias, lymphedema, and hematoma [11-13]. Moreover, for many patients, SLNB is more morbid than partial mastectomy and therefore axillary surgery is not considered therapeutic in early-stage breast cancer [14,15]. In this context, there is a need to evaluate other less invasive and preferably cheaper diagnostic tools that could replace SLNB without compromising patient care.

The Clinical Decision Support Systems (CDSSs) are a data elaboration system able to extract information available from retrospective cases for supporting clinical decision-making improving its quality [16,17]. Our hypothesis is that CDSSs can be developed to predict the sentinel lymph node positivity. Specifically, if a high prediction accuracy is reached, both intervention of biopsy could be avoided and the efficiency of surgical intervention could be reduced, in terms of time and costs.

Some approaches were proposed in the literature that use histological [18], genetic [19] or radiomic data to predict the status of the lymph node involvement with performing results [20-23]. Among the different proposals in the literature Cancer Math (CM) [24] is a general-purpose online software aiming at the prediction of patient survival for some kinds of tumours. In the breast cancer case, a probability estimation for lymph nodes positive status is included to take into account non-lethal metastatic diffusion processes, which is not further singularly inspected according to its predictive properties [25,26].

In this work we present the preliminary results of validation analysis aimed to evaluate the usefulness of CancerMath (CM) web-calculator in clinical practice for predicting the status of sentinel lymph nodes in patients with T1 breast cancer and clinically negative nodes. We first considered the classification performance of the CM algorithm on a sample of patients visiting our Institute with breast carcinoma. Subsequently, we trained the algorithm and evaluated the classification results obtained both by considering the same set of prognostic factors used by the online software, and then by adding other known prognostic factors such as the proliferation marker Ki67 and the epidermal growth factor receptor 2 (HER2).

**Methods**

**Experimental data**

The cohort of patients considered in our study was negative following clinical and radiological examinations carried out in the preoperative phase [4] and having undergone the OSNA procedure, which is characterized by a small tumour diameter in the T1 category, thus taking into account sizes in the range 1-20 mm, further subdivided into T1a (1-5 mm), T1b (5-10 mm) and T1c (10-20 mm). The overall number of patients in the dataset was 595, with 98 positive cases, recorded from 2016-2018 in Istituto Tumori “Giovanni Paolo II” of Bari, Italy. In addition to the age at diagnosis, in the analysis we included the histological outcome as described below. This retrospective observational study was approved by our Institute’s Scientific Board.

**Histological evaluation procedure**

According to the European guidelines in case of uncertain or suspect clinical or radiological examination of the axillary status, the surgeon decides either to perform lymph nodes axillary dissection during surgery or, in absence of suspicion, to proceed with the sentinel lymph nodes biopsy evaluation. The OSNA method is an isothermal procedure which employs a nucleic acids rapid amplification technology in order to detect the mRNA expression level in the cytokeratin 19, an epithelial cells marker not normally present in lymph nodes. The OSNA system is in compliance with the European directive about the in vitro diagnostic 98/79/CE (CE-IVD), and it is therefore approved for diagnostic purposes throughout Europe.

Ultrasound devices allow for the implementation of multiple 14-16 G core biopsy sampling, thus yielding the collection of histological samples. The pathological anatomy department obtains a material with formalin containing picked whips. Subspecialty departments of breast disease in our Institute are involved in immunohistochemical analysis required to yield prognostic factors associated with histological type, as well as percentage values for estrogen receptor (ER), progesterone receptor (PgR), human epidermal growth factor receptor 2 (HER2) and Ki-67 proliferation marker. The Elston-Elis modification of the Scarff-Bloom-Richardson grading system yields a three-scale tumor grade (G). The latter is
a concise measure, based on duct structures, of mitotic rate, size and configuration of tumour cells nuclei, which reads as low (G1), intermediate (G2) and high (G3) [27]. The immunohistochemical expression follows St. Gallen subtypes: luminal A (ER+ and/or PgR+, HER2-, low Ki67), luminal B (ER+ and/or PgR+, high Ki67 for HER2+/any Ki67 for HER2-), HER2 positive (ER/PgR-, HER2+) and triple negative (ER/PgR-, HER2-).

**Prediction models**

The CancerMath (CM) algorithm implements an online free platform [24] and with an open source code aims at modelling the diffusion of cancer cells belonging to the primary lesion. CM estimates the probability of lymph node involvement on the basis of prognostic factors such as cancer mass diameter measurement, age, histological type, grading, and presence of both estrogen ER and PgRs. The latter online version is characterized by some preset parameters yielded by a training procedure executed on a widespread database concerning multiple clinical conditions. The online code does not include any training, but it is possible to implement it according to published instructions, such that to adapt the algorithm in a more focused perspective.

Probability of positive sentinel lymph nodes is estimated by an exponential mathematical model [24]:

\[ L_n = 1 - \exp\left(-Q_n D \prod_{i=1}^{5} g_i\right), \] (1)

where D is the diameter of the primary cancer mass, \( g_i \) are the values of the 5 variable parameters age, grading, histological type, ER (Pos/Neg), PgR (Pos/Neg), while \( Q_n \) is an interpolation parameter referred to the whole population. The values of both \( Q_n \) and the parameters \( g_i \) are determined using a training procedure on the training set which produces the measure of the prognostic factor correspondent impact as a statistically independent cause [25,26]. The training procedure initially determines the \( Q_n \) value for the entire population, setting in this phase the parameters \( g_i \) all equal to 1. Then, the recently determined value of \( Q_n \) is set in each sub-population and the same procedure is carried out for each range of values of the prognostic factors.

Since the grading is a missing information for some patients of our dataset, as well as the histological type, we trained the model used by CM by including factors related to them, that is Ki67 [28,29] and HER2 [30,31], respectively. Therefore, starting from CM algorithm we investigated three other classifiers, the first one obtained by training the same model on our dataset (A), and the other ones obtained by adding the proliferation marker Ki67 (Pos/Neg) (B) or either HER2 (Pos/Neg) (C).

**Performance evaluation**

The CM performances computed using the online software were compared with the ones obtained by retraining the classification algorithm CM on our dataset with the same features (classifier A) and then adding separately the two prognostic factors Ki67 (classifier C) and HER2 (classifier B).

The classification performances of CM on line application and A, B, C, models both on Hold-out training and test set are evaluated in terms of Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve and, once identified the optimal threshold by Youden’s index on ROC curves [32], we have also calculated:

- **Accuracy** = \((TP + TN)/(TP + TN + FP + FN)\)
- **Sensitivity** = \(TP/(TP + FN)\)
- **Specificity** = \(TN/(TN + FP)\)

where TP and TN stand for true positive (number of cases with positive sentinel lymph correctly classified) and true negative (number of cases with negative sen-

---

**Table 1. Number of patients with positive lymph nodes according to considered prognostic factors**

<table>
<thead>
<tr>
<th></th>
<th>N. patients</th>
<th>N. positive</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>595</td>
<td>98</td>
<td>ER</td>
<td>Patients, n</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>negative</td>
<td>74</td>
</tr>
<tr>
<td>Age-years</td>
<td></td>
<td></td>
<td>positive</td>
<td>521</td>
</tr>
<tr>
<td>31-40</td>
<td>32</td>
<td>6</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>41-50</td>
<td>159</td>
<td>44</td>
<td></td>
<td>462</td>
</tr>
<tr>
<td>51-60</td>
<td>191</td>
<td>51</td>
<td>negative</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>positive</td>
<td>462</td>
</tr>
<tr>
<td>61-70</td>
<td>139</td>
<td>14</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td>71-80</td>
<td>71</td>
<td>3</td>
<td>Ki67</td>
<td>426</td>
</tr>
<tr>
<td>81-90</td>
<td>3</td>
<td>0</td>
<td>negative</td>
<td>68</td>
</tr>
<tr>
<td>Diameter</td>
<td></td>
<td></td>
<td>positive</td>
<td>169</td>
</tr>
<tr>
<td>T1a</td>
<td>18</td>
<td>2</td>
<td>HER2</td>
<td>30</td>
</tr>
<tr>
<td>T1b</td>
<td>160</td>
<td>19</td>
<td>negative</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>positive</td>
<td>527</td>
</tr>
<tr>
<td>T1c</td>
<td>417</td>
<td>77</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>Grading</td>
<td></td>
<td></td>
<td>unknown</td>
<td>4</td>
</tr>
<tr>
<td>G1</td>
<td>41</td>
<td>14</td>
<td>Histologic type</td>
<td>1</td>
</tr>
<tr>
<td>G2</td>
<td>73</td>
<td>14</td>
<td>ductal</td>
<td>399</td>
</tr>
<tr>
<td>G3</td>
<td>36</td>
<td>9</td>
<td>lobular</td>
<td>39</td>
</tr>
<tr>
<td>Unknown</td>
<td>445</td>
<td>61</td>
<td>unknown</td>
<td>157</td>
</tr>
</tbody>
</table>

JBUON 2021; 26(3): 722
CancerMath predicts nodal metastasis in breast cancer

JBUON 2021; 26(3): 723

Tinel lymph correctly classified, while FP (number of negative cases identified as positive) and FN (number of positive cases identified as negative) are false positive and false negative ones, respectively.

Specifically, the prediction performances of A, B and C models obtained on 100 ten-fold cross-validation rounds and summarized in terms of median, 1st and 3rd quartile.

**Results**

**Datasets**

Patient characteristics are summarized in Table 1. A total of 595 patients aged between 31 and 83 years (with a median, first and third quartile of 55, 48 and 66 years, respectively) were included in the study of which 98 of these were positive on histological examination of the sentinel lymph nodes.

The variables considered by CM models were diameter of the primary cancer mass, weighted by the other variables corresponding to age, histological type, grading and percentage values of ER, PgR, Ki67, and HER2. In this work we included Ki67 proliferation marker among prognostic factors and HER2. Since the grading is a missing information for some patients of our dataset, as well as the histological type, we trained the model used by CM by including factors related to them, that is Ki67 and HER2, respectively (p-values chi-square test <0.05) [29-31].

**Performances comparison**

The prediction of sentinel lymph node status obtained using CM web-calculator for our dataset was not highly performing. In fact, it reached an AUC median value of 61.8% and in particular median accuracy, sensitivity and specificity values of 57.5%, 61.2% and 56.7%, respectively.

Then we trained the model on our dataset to evaluate the possible increase in performance with the same features set. Figure 1 shows the ROC

| Table 2. Classification performances obtained by online CM application and by training A, B and C models on our dataset. The prediction performance of A, B and C models were evaluated on 100 ten-fold cross-validation rounds and summarized in terms of median, 1<sup>st</sup> and 3<sup>rd</sup> quartile |
|-----------------|---------|---------------------|---------------------|---------------------|---------------------|
|                | Percent | Performances measure |
| CM on line     | AUC     | 61.8                |
|                | Acc     | 57.5                |
|                | Sens    | 61.2                |
|                | Spec    | 56.7                |
| (A) CM features| AUC     | 69.4 (68.9-69.7)    |
|                | Acc     | 61.5 (58.7-65.3)    |
|                | Sens    | 71.4 (65.3-75.5)    |
|                | Spec    | 58.9 (55.5-64.9)    |
| (B) CM features +Her2| AUC | 68.9 (68.4-69.2) |
|                | Acc     | 60.1 (58.5-62.4)    |
|                | Sens    | 73.0 (67.3-75.5)    |
|                | Spec    | 57.8 (55.3-61.6)    |
| (C)CM features +ki67| AUC | 68.8 (68.4-69.3) |
|                | Acc     | 60.8 (58.8-65.4)    |
|                | Sens    | 70.4 (67.3-74.5)    |
|                | Spec    | 58.9 (55.5-62.6)    |

Figure 1. ROC curves obtained trained CM algorithm on our dataset and evaluated on 100 ten-fold cross-validation rounds.
CancerMath predicts nodal metastasis in breast cancer

724

JBUON 2021; 26(3): 724

Curves obtained on 100 ten-fold cross-validation rounds which were stable around an AUC median value 69.4%.

Table 2 summarizes the performances statistics obtained on 100 ten-fold cross-validation rounds. Specifically, classification performances of CM trained on our dataset in cross-validation show a significant gain in sensitivity with respect to the online version, still maintaining the same level for specificity (Student-t test p<0.05). Indeed, CM web-calculator reached a specificity of 57% and a sensitivity of 61%. By training the same algorithm in cross-validation with the same features the sensitivity increased to 71.4% and specificity dropped to 58.9%.

Since the grading is a missing information for some patients of our dataset, as well as the histological type, we trained the model used by CM by including factors related to them, that is Ki67 and HER2, respectively (x^2 test, <0.05). The entering of Ki67 or HER2 did not significantly improve the general accuracies with respect the CM model trained on our dataset with the same features used by web-calculator. However, by introducing HER2 in the CM model, we reached a median sensitivity value of 73%, without losing specificity.

Obtained performances by the models analyzed in sub-samples stratification determined by means of specific values or intervals that usually typify a clinical condition [33-37] are summarized in Table 3. Specifically, they regard three age classes for age >60, triple-negative, even if the latter feature categories were associated with small sub-samples. When HER2 was introduced, there was an increase in sensitivity in the prediction of sentinel lymph node positivity of grade 3 and luminal B tumors, while the introduction of Ki67 affected the sensitivity of elderly patients (age>60) and grade 2 tumors.

Discussion

OSNA technique is time-consuming and expensive, but it is still the intraoperative exam with the currently best results. Nevertheless, SLNB is an invasive surgical procedure and although it has fewer complications and is less severe than axillary lymph node dissection, it is not a risk-free procedure and is considered non-therapeutic in early stage breast cancer [38,39]. In this work we evaluated the usefulness of clinical decision support system to predict the lymph node status in T1 breast cancer as alternative procedure to SLNB.

The online CM algorithm studied in this work aimed to estimate the probability of having positive lymph nodes of breast cancer although this probability computation is included in a broader framework regarding the oncological patient survival analysis [40,41]. Compared to other online software, CM shows higher performance as reported in several validation studies [42-45].
targeted data, thus avoiding the bias caused by a general purpose software trained according to a less focused dataset. The introduction of new factors HER2 and Ki67 affects only some categories of sub-samples [46].

The CM performances have been validated in different studies [40-42], but only in one study [41] the authors consider the package related to the lymph node involvement probability computation on a sample of the population of the South-Eastern Asia. In this work the authors highlight that the median probability estimated by CM of having a lymph node involvement is equal to 40.6%, underestimating the real one of the samples under study (43.6%). However, the authors did not validate the model by measuring the classification performance of the lymph node status on the studied sample and not trained the model on their dataset.

Although the training of the model on the sample of T1 patients has brought a significant improvement in performance, the general performance does not yet allow a clinical application of the algorithm. However, the experimental results encourage future developments aimed at introducing features of a different nature in the CM model.

However, it could be possible to improve classification performances also in terms of specificity by extending the features set to include multifocality, lymphovascular invasion as demonstrated in some state-of-the-art works relating to the broader problem of predicting the lymph node status [40,41]. Improvement is achieved also through radiomics approach: the enriched information coming from imaging setups, like diffusion wavelet and dynamic contrast-enhanced magnetic resonance imaging (DWI and DCE-MRI), is exploited using logistic regression models [20], convolution neural networks [21] and least absolute shrinkage and selecting operator [22,23]. The AUC index in these studies is always in the range 75-85%. Moreover, most works at the state-of-art propose non-sentinel lymph node status predictive models by using nomograms of clinical and pathologic variables [47-51]. On the other hand, the literature is poor in studies aimed to the development of sentinel lymph nodes predictive models studies that can help avoid the SLNB procedure, yet expensive and not free from complications [18-19,52]. In a study [52] the authors proposed a prediction model for sentinel lymph node metastasis using genetic features, tumor size and lymphovascular invasion in ER-positive and HER2-negative (ER+/HER2−) breast cancers, thus reaching an AUC value of 0.883. In another study [53] the authors developed a prediction model for detecting the negativity of sentinel lymph node in order to reduce additional axillary lymph node procedure, by reducing time and costs of operations, and also unnecessary axillary dissections for T1 breast cancers.

Conclusions

The online CM software is trained using a large sample of patients with heterogeneous characteristics. Our study underlines the variation in performances given by a specific training regarding more focused databases on the targeted early stage lymph node metastasis. Even if local data miss required information about a prognostic factor, it can be substituted with a correlated one. In our case, the introduction of HER2 prognostic factor yielded further improvement in the selection of true positive cases.

Harmonizing the histological data with the ones obtained from the other information sources will be the aim of our future works to increase prediction accuracy, such as multifocality, lymphovascular invasion and tumor location [54], genetics [19] and above all radiomics [55,56]. Finally, other more sophisticated machine learning models can be implemented and evaluated to optimize classification performances [57,58]. By achieving high levels of accuracy, the use of such support system would make it possible to avoid both the sentinel lymph node procedure, by reducing time and costs of operations, and also unnecessary axillary dissections for T1 breast cancers.

Author contributions


Acknowledgements

This work was supported by funding from the Italian Ministry of Health “Ricerca Corrente 2018–2020”.

Conflict of interests

The authors declare no conflict of interests.
References


