Validating the MSKCC nomogram and a clinical decision rule in the prediction of non-sentinel node metastases in a Portuguese population of breast cancer patients

José Luís Fougo, Francisco Senhorães Senra, Cláudia Araújo, Teresa Dias, Mariana Afonso, Conceição Leal, Mário Dinis-Ribeiro

**Abstract**

**Background:** In order to assess the risk of non-sentinel node involvement in breast cancer patients, some prediction tools have been developed and tested. However, a wide range of results are observed. We tested a simplified clinical decision rule, against the complex nomogram created from the MSKCC sentinel node database.

**Methods:** Two single institutional consecutive series of patients with a positive SN, submitted to SN biopsy plus axillary dissection from June 1999 to October 2007, were evaluated. A receiver operating curve was drawn and the area under the curve was calculated as well as the negative predictive value for both tests, assuming discriminative values of 10 and 15%.

**Results:** Considering the derivation series, our results showed an area under the curve of 0.69 for both our clinical decision rule and the MSKCC nomogram. The analysis of the validation series showed an area under the curve of 0.65 for our clinical decision rule and of 0.67 for the MSKCC nomogram. The nomogram results are inferior to those found in the original population and are similar to our clinical decision rule results.

**Conclusions:** Individual centres should develop and prospectively test their own clinical decision rules, based on their institutional Sentinel Node data.

**Introduction**

During the last decades, the trend in the treatment of breast cancer (BC) patients is to conserve, both the breast and axillary nodes. Breast cancer conserving treatment (local excision of the tumour plus radiotherapy) has proven to be equivalent to total mastectomy in terms of distant metastases and overall survival.1,2 The last decade consistently established a way to assess the nodal status of a BC, sparing the patients from the morbidity of a formal axillary lymph node dissection (AD): the Sentinel Node (SN) concept.3 Complete AD was reserved for those patients with a positive SN. The Portuguese Society of Senology and other Cancer Societies agree with this attitude.4,5

The analysis of SN series has been showing that 50–60% of SN positive BC patients do not have any other positive nodes in the AD specimen.6,7 Therefore, in those patients, AD could be considered excessive treatment. In the last few years, researchers have been trying to relate patient, tumour and SN characteristics to the presence of non-sentinel node (NSN) metastases.8–12 Some predictive models have been developed.13–16 Among those, the Memorial Sloan Kettering Cancer Center (MSKCC) nomogram gave rise to a worldwide reaction, being tested in different settings.17–23 The results of its application vary between series, and suggestions to its adaptation have been made, in order to render it a more valuable tool to counsel SN positive patients.24,25

Previously we have derived a simple clinical decision rule (CDR) that includes the size of the tumour (pT1 vs pT2/3), multifocality (absence vs presence) and lymph–vascular invasion (absence vs presence).16

This study assessed two different sets of patients: the group of 143 patients from which the CDR was derived and a new group of 98 patients. The aim of the study was to validate our CDR and to test it against the MSKCC nomogram.
Patients and methods

Patient population

The two groups of patients studied in this analysis were composed of 143 (derivation group) and 98 (validation group) SN positive, AD submitted patients, operated from June 1999 to October 2007. All patients were consecutive. All of them participated in institutional review board approved protocols for the application of SN concept to BC patients, at the Portuguese Institute of Oncology Francisco Gentil, Porto Centre (PIO-P), a tertiary cancer centre.

Clinical and pathology data from the patient, the tumour, sentinel and non-sentinel nodes were prospectively gathered in an institutional database.

Preoperative axillary ultrasound was never done and none of the patients underwent neo-adjuvant systemic treatment.

Surgical procedures

We used a combined, triple, technique to identify the SN. This technique includes peritumoural injection of a radioactive colloid, the execution of a preoperative lymphoscintigraphy, the injection of a vital dye in the subareolar plexus and the use of a hand-held gamma-probe. A blue, hot or hot-and-blue node was considered an SN; any other suspicious node was also considered an SN.

When the SN was positive we performed an AD (Berg levels I + II). The axillary vein, the thoraco-dorsal neuro-vascular bundle and the long thoracic nerve were identified, dissected and spared; the costo-brachial sensitive nerves are usually not spared.

Pathology workup

Perioperative examination was done on 104 patients (derivation group, 72.7%) and on 98 patients (validation group, 100%). Each SN was cut longitudinally into slices of approximately 2 mm. Imprint cytology was made from all cut surfaces and the most suspicious slice was frozen for immediate examination. Definitive pathological analysis consisted in haematoxylin–eosin section. Immunohistochemistry was never done.

Non-sentinel nodes smaller than 5 mm were fixed and embedded unsliced; nodes with a major axis greater than 5 mm were bisected and one slice was fixed and embedded. One haematoxylin–eosin section was performed on each paraffin block.

The tumours were classified according to the World Health Organisation Histological Classification of Breast Tumours and graded according to the Nottingham Histological Grading System.

The MSKCC nomogram analysis

A nomogram, also known as a slide rule, is a statistical tool, a two-dimensional diagram designed to allow the approximate graphical computation of a function.

The MSKCC nomogram was developed and described by van Zee et al.13 It was downloaded from the MSKCC website (http://www.mskcc.org/mskcc/15938.cfm) to a Personal Digital Assistant (Palm Tungsten T3™).

Routine frozen section was only done on 202 patients. Pathological size was defined by a microscope rule. Nuclear grade was estimated in all patients; histological mixed and non–ductal/nonlobular tumours were grouped as ductal type.20 Method of detection chosen was “serial H&E” for all patients. Multifocality was considered either clinical or microscopic. Oestrogen receptors were considered positive if more than 10% of the tumour cells showed immunohistochemical reaction.

Predicted and observed values were calculated for both cut-off levels of 10% and 15%. For example, a result of the MSKCC nomogram of 23% was assumed as an NSN positive predicted result and compared to its observed result; a result of the MSKCC nomogram of 8% was assumed as an NSN negative predicted result and compared to its observed result as well.

The cut-off level of 15% was chosen because it was the classical lymph node risk of involvement cut-off at which the proposal for prophylactic lymph node dissection was made. The cut-off level of 10% was chosen as suggested by the works of Lambert,20 Park,26 Coutant27,28 or Cserni29 and because it was closer to the previously defined risk of 11% with our CDR (see below).

Table 1

Clinical and Pathological characteristics of the two groups of patients.

<table>
<thead>
<tr>
<th></th>
<th>PIO-P Derivation (n = 143)</th>
<th>Validation (n = 98)</th>
<th>MSKCC (n = 373)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>≤50</td>
<td>56 (39.2)</td>
<td>35 (35.7)</td>
<td>157 (42.1)</td>
<td></td>
</tr>
<tr>
<td>&gt;50</td>
<td>87 (60.8)</td>
<td>63 (64.3)</td>
<td>216 (57.9)</td>
<td></td>
</tr>
<tr>
<td>Pathologic tumour size (mm)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≤5</td>
<td>0 (–)</td>
<td>2 (2.0)</td>
<td>13 (3.5)</td>
<td></td>
</tr>
<tr>
<td>6–10</td>
<td>4 (2.8)</td>
<td>6 (6.1)</td>
<td>49 (13.1)</td>
<td></td>
</tr>
<tr>
<td>11–20</td>
<td>59 (41.3)</td>
<td>40 (40.8)</td>
<td>166 (44.5)</td>
<td></td>
</tr>
<tr>
<td>21–30</td>
<td>51 (35.7)</td>
<td>30 (30.6)</td>
<td>93 (24.9)</td>
<td></td>
</tr>
<tr>
<td>31–50</td>
<td>27 (18.9)</td>
<td>16 (16.3)</td>
<td>41 (11.0)</td>
<td></td>
</tr>
<tr>
<td>≥51</td>
<td>2 (1.4)</td>
<td>4 (4.1)</td>
<td>11 (2.9)</td>
<td></td>
</tr>
<tr>
<td>Tumour type and nuclear grade</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ductal, I</td>
<td>3 (2.1)</td>
<td>15 (15.3)</td>
<td>11 (2.9)</td>
<td></td>
</tr>
<tr>
<td>Ductal, II</td>
<td>33 (23.1)</td>
<td>55 (56.1)</td>
<td>175 (46.9)</td>
<td></td>
</tr>
<tr>
<td>Ductal, III</td>
<td>106 (74.1)</td>
<td>22 (22.5)</td>
<td>129 (34.6)</td>
<td></td>
</tr>
<tr>
<td>Lobular</td>
<td>1 (0.7)</td>
<td>6 (6.1)</td>
<td>58 (15.5)</td>
<td></td>
</tr>
<tr>
<td>Lymph–vascular invasion</td>
<td></td>
<td></td>
<td></td>
<td>0.018</td>
</tr>
<tr>
<td>No</td>
<td>67 (46.8)</td>
<td>68 (69.4)</td>
<td>219 (58.7)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>76 (53.2)</td>
<td>30 (30.6)</td>
<td>154 (41.3)</td>
<td></td>
</tr>
<tr>
<td>Multifocality</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No</td>
<td>119 (83.2)</td>
<td>79 (80.6)</td>
<td>241 (64.6)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>24 (16.8)</td>
<td>19 (19.4)</td>
<td>132 (35.4)</td>
<td></td>
</tr>
</tbody>
</table>
The PIO-P decision rule analysis

Our decision rule was based on the following tumour related variables: microscopic tumour size (pT1 vs pT2-3), multifocality (absence vs presence) and lymph–vascular invasion (absence vs presence). When none of those variables were present (that is, in view of a unifocal, pT1 tumour without lymph–vascular invasion) the risk of NSN positivity was inferior to 11%.16

If a case presented itself without any of the above defined characteristics, it was assumed as an NSN negative predicted result and compared with its observed result; if a case presented itself with any of those variables, it was assumed as an NSN positive predicted result and compared to the observed result.

Statistical analysis

A prospective, computerized, database on SN procedures was undertaken from June 1999 (SPSS for Windows, version 17, SPSS Inc., Chicago, Illinois, USA).

The predicted probability of NSN positivity was calculated for each patient by using the MSKCC nomogram and our CDR, in both derivation and validation groups.

For each group of patients, a receiver operating characteristic curve (ROC) was drawn, and the area under the curve (AUC) was calculated to assess the discriminative power of the nomogram. A calibration plot was drawn showing the actual versus the nomogram-predicted probability.

All reported P values were two-sided.

Results

The clinical and pathological characteristics of the two groups of patients are listed in Tables 1 and 2. The median age was 54 years (range: 20–78 years) and 57.5 years (range: 27–80 years), for the derivation and validation groups, respectively. The median pathological tumour size was 22 mm (range: 7–80 mm) and 21 mm (3–80 mm). The mean number of excised SN was 1.9 (±1.1) and 1.8 (±0.9). The mean number of excised nodes at AD was 13.9 (±5.1) and 14.6 (±5.1).

Non-sentinel node metastases were found in 66 (46.2%) and in 43 patients (43.9%).

Table 3 shows the predicted and observed results for the MSKCC nomogram, for both cut-off points at 10 and 15% and for the PIO-P clinical decision rule, for a cut-off point at 11%.16

### Table 2
Clinical and pathological characteristics of the two groups of patients.

<table>
<thead>
<tr>
<th></th>
<th>PIO-P</th>
<th>MSKCC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Derivation (n = 143)</td>
<td>Validation (n = 98)</td>
</tr>
</tbody>
</table>

### Table 3
Predicted and observed results for the MSKCC nomogram and PIO-P CDR.

<table>
<thead>
<tr>
<th></th>
<th>True + n</th>
<th>True - n</th>
<th>False + n</th>
<th>False - n</th>
<th>Se% (95%CI)</th>
<th>Sp% (95%CI)</th>
<th>LR + % (95%CI)</th>
<th>LR - % (95%CI)</th>
<th>PPV% (95%CI)</th>
<th>NPV% (95%CI)</th>
</tr>
</thead>
</table>
Application of the MSKCC nomogram and the PIO-P CDR to the PIO-P derivation group of patients

We applied the MSKCC nomogram to our sample of 143 patients. This model predicted the likelihood of NSN metastases as an AUC of 0.691 (95% confidence interval 0.604–0.778). We also determined the AUC when applying our CDR, obtaining a result of 0.695 (95% confidence interval 0.608–0.781) (Fig. 1A). The calibration plot is shown in Fig. 2A.

Application of the MSKCC nomogram and the PIO-P CDR to the PIO-P validation group of patients

The MSKCC nomogram, applied to the validation group, predicted the likelihood of NSN metastases as an AUC of 0.673 (95% confidence interval 0.554–0.871); our CDR issued an AUC of 0.654 (95% confidence interval 0.533–0.775). The two curves are superimposed (Fig. 1B). The calibration plot is shown in Fig. 2B.

---

**Fig. 1.** ROC curves of the MSKCC nomogram and of the PIO-P CDR, for the derivation group (A) and for the validation group (B).

**Fig. 2.** Calibration plots of the MSKCC nomogram and of the PIO-P CDR, for the derivation group (A) and for the validation group (B).
Discussion

“Prediction is very difficult, especially about the future.” This quote from Niels Bohr, an XX century Danish physicist, is still applicable today, concerning the prediction of NSN involvement in SN positive BC patients.

Several reports have been published with reference to this issue, defining variables associated to NSN positivity and promoting scoring systems/nomograms; these scoring systems intend to quantify the risk for the presence of NSN metastases, hence helping doctors and patients to decide when or not to do an AD.

Our CDR, when applied to the validation series, obtained ROC/AUC results similar to those obtained with the derivation series, with acceptable Sensitivity (91%) and Negative Predictive Value (NPV, 82%) rates. This sensitivity level is within the suitable limits for the general application of the Sentinel Node concept, that is, between 90 and 95%. This means that the application of this CDR to positive SN BC patients will not have a false-negative result risk markedly different from the false-negative risk of applying the SN concept to a clinically node-negative BC patient.

The derivation group yielded an AUC result of 0.695 and the validation group a result of 0.654, belonging to the poor discrimination interval (0.61–0.70).30

Most groups that tested the MSKCC nomogram, or their own predictive methods, published results within the spectrum of fair discrimination (0.71–0.80).18–20,22,28,31–33 but some reported poor discrimination results as well25,34–37 (Table 4).

Nevertheless, as already stated,38,39 predictive tools are better than expert opinion.

International, widespread, application of a single predictive method should be questioned, on the basis of the variable results of the validation. Likewise, there is a worldwide profound variability on the interpretation of the SN concept, on the SN identification technique and on the SN pathology workup. In this context, Klar et al.40 recently stated that the standardization of the surgical procedure and of the pathological assessment of the SN would be desirable. However, this may be very difficult to obtain.

Moreover, as declared by Degnim et al.,24 the accuracy of the prediction by a model can degrade as the model is transported from one population to another.

These views should prompt breast surgeons and their centres to develop their own nomograms/scoring systems/clinical decision rules, based on their SN experiences and settings3,35 and avoiding the use of SN related variables in those systems.16

Our CDR is characterized by depending only upon three dichotomous tumour variables, dispensing the use of sentinel node features. Additionally, the status of these three features can be defined preoperatively.

Sentinel node metastases size has been reported as one of the most influencing variables for the prediction of NSN positivity, that is, the bigger the SN metastases, the higher the risk for SN involvement. This seems logical; nevertheless, there are numerous reports of the prevalence of NSN metastases in the presence of SN micrometastases12,41 or isolated tumour cells (ITC).42 This fact should remind surgeons about the classical pitfalls of SN concept and its clinical application: a heavily involved SN could redirect the breast lymph drainage to another node, and this second echelon node could be biopsied and wrongly interpreted as the SN. So, “our” SN could be a wrongly or the least informative SN.

Besides that, the MSKCC nomogram results have been difficult to interpret in the presence of SN micrometastases.

Studies from Alran22 and Gur33 have shown limited results in the use of the MSKCC nomogram when the SN had only small amounts of tumour (≤2 mm).

In the same way, Cserni and collaborators44 highlighted that diverse interpretation of the TNM pathology regulations to define micrometastasis or ITC may imply deep fluctuation among the prediction of NSN involvement.

So, most authors still recommend the AD completion when the SN shows micrometastasis or ITC.42,45,46

A “big” SN metastasis may be useful to assist in the decision to perform a complementary AD, but a “small” SN metastasis could not be useful to define a negative predictive value.

One should admit that, in future, preoperative axillary ultrasound will be a routine exam for BC patients. Moreover, it may be possible to identify and to biopsy with a needle the SN.47–49 So, the decision to forego AD could not be based on SN characteristics, because the SN will not be available to be fully analysed.

Another way of assessing the usefulness of the scoring systems is its Clinical Interest. This parameter was stressed by Coutant27 and relies on the sensitivity, specificity, NPV and Positive Predictive Value (PPV) of the method, balanced with its ability to recruit a sufficient number of patients to be spared from AD completion.

The GCFFC (Surgeons Association of the French Federation Against Cancer) paper41 defined a predictive model based on four characteristics; this model obtained an AUC of 0.66 (FN, PPV and NPV were not shown) and defined a group of 29.8% of the patients with an NSN involvement risk of <10%.

A recently published multicentric, prospective, French work,38 which compares the results of nine predictive models, points out that three of the models revealed a false-negative rate superior to 10% and stressed that the target for the FN rate should be 5%, resembling the axillary dissection FN rate. We can add that it resembles the FN rate for the SN concept (5–10%). This same work found that a scoring system, known as the Tenon Score,15 was able to select 48% of the studied patients as low risk patients, which represents an excellent clinical interest.

Our model issued an FN rate of 9% and was able to select 18.2 (23 – 3 out of 143) and 22.5% (18 / 4 out of 98) of the patients, respectively from the derivation and from the validation groups. This means that 1 out of 5 positive SN patients can be spared from AD completion with the use of our CDR; this is, in our opinion, of considerable clinical interest, meaning a safe, reliable, effort to avoid potentially unnecessary surgical procedures.

As was stressed by Coutant and collaborators,50 predictive models have to be simple to use and have an excellent NPV. Our model will be tested prospectively, in order to assess its capability of preoperative discrimination, namely before tumour excision and SN biopsy.

Table 4

<table>
<thead>
<tr>
<th>Author</th>
<th>Model evaluated</th>
<th>n</th>
<th>AUC result</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present work, 2009</td>
<td>MSKCC</td>
<td>58</td>
<td>0.673</td>
<td>0.554–0.871</td>
</tr>
<tr>
<td>CDR</td>
<td>98</td>
<td>0.654</td>
<td>0.531–0.775</td>
<td></td>
</tr>
<tr>
<td>Sony et al., 2005</td>
<td>MSKCC</td>
<td>149</td>
<td>0.75</td>
<td>0.668–0.814</td>
</tr>
<tr>
<td>Smitd et al., 2005</td>
<td>MSKCC</td>
<td>222</td>
<td>0.76</td>
<td>0.69–0.83</td>
</tr>
<tr>
<td>Lambert et al., 2006</td>
<td>MSKCC</td>
<td>200</td>
<td>0.71</td>
<td>Not shown</td>
</tr>
<tr>
<td>Alran et al., 2007</td>
<td>MSKCC</td>
<td>588</td>
<td>0.724</td>
<td>0.677–0.771</td>
</tr>
<tr>
<td>Pal et al., 2008</td>
<td>MSKCC</td>
<td>118</td>
<td>0.68</td>
<td>Not shown</td>
</tr>
<tr>
<td>Coursant et al., 2009</td>
<td>MSKCC</td>
<td>561</td>
<td>0.78</td>
<td>0.76–0.81</td>
</tr>
<tr>
<td>Cho et al., 2008</td>
<td>MSKCC</td>
<td>82</td>
<td>0.786</td>
<td>0.687–0.884</td>
</tr>
<tr>
<td>Winston et al., 2006</td>
<td>MSKCC</td>
<td>82</td>
<td>0.825</td>
<td>0.732–0.917</td>
</tr>
<tr>
<td>van la Parra et al., 2009</td>
<td>MSKCC</td>
<td>182</td>
<td>0.712</td>
<td>0.635–0.790</td>
</tr>
<tr>
<td>Zagaj et al., 2007</td>
<td>MSKCC</td>
<td>276</td>
<td>0.722</td>
<td>Not shown</td>
</tr>
<tr>
<td>Kohut et al., 2008</td>
<td>MSKCC</td>
<td>77</td>
<td>0.62</td>
<td>0.55–0.68</td>
</tr>
<tr>
<td>Stanford</td>
<td>MSKCC</td>
<td>77</td>
<td>0.74</td>
<td>0.67–0.80</td>
</tr>
<tr>
<td>COUFL et al., 2009</td>
<td>MSKCC</td>
<td>383</td>
<td>0.74</td>
<td>0.68–0.80</td>
</tr>
<tr>
<td>Dauphine et al., 2007</td>
<td>MSKCC</td>
<td>39</td>
<td>0.63</td>
<td>0.46–0.81</td>
</tr>
<tr>
<td>Poirier et al., 2007</td>
<td>MSKCC</td>
<td>209</td>
<td>0.687</td>
<td>Not shown</td>
</tr>
</tbody>
</table>
Authorship
Conception and design – JLF, MDR
Data acquisition – JLF, FSS, TD
Pathologic review – MA, CL
Analysis and interpretation of data – JLF
Manuscript draft – JLF
Statistical data revision and analysis – JLF, MDR
Bibliographic research – JLF
Critical revision – JLF, MA, CL, CA, MDR
Supervision – MDR

Conflict of interest statement
None declared.

Appendix. Supplementary data

Supplementary data related to this article can be found online at doi:10.1016/j.breast.2010.10.009.

References


