Methodology and implications of lymphatic mapping and sentinel lymphadenectomy
Tanis, P.J.

Citation for published version (APA):
Tanis, P. J. (2002). Methodology and implications of lymphatic mapping and sentinel lymphadenectomy

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER FOURTEEN

The Illusion of the Learning Phase for Lymphatic Mapping

Departments of Surgery and Radiotherapy, The Netherlands Cancer Institute / Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands

Ann Surg Oncol, in press
The sentinel node is the first lymph node reached by metastasising cells from a primary tumour. The sentinel node concept has been proven to be valid in numerous studies.\(^1\) False-negative biopsies can occur because of re-routing of lymphatic drainage by tumour blockage, variability of lymph flow, sampling error by the pathologist and inexperience of the surgeon.\(^3\) The latter point has been addressed in the literature several times.\(^4,5\) One of the main topics in this discussion concerns the learning phase. There are two aspects to identify the sentinel node: “finding” the node and establishing that the excised lymph node is indeed the “correct” node. The ability to find the node is expressed as the identification rate. The sensitivity reflects the ability of the surgeon to remove the correct node. The latter is indicated by the absence of tumour-positive nodes elsewhere in the regional basin in case of a tumour-negative sentinel node. Both elements play a role in the learning phase.

Some data about the initial experiences in sentinel node biopsy of individual surgeons have been published. Eleven surgeons participated in the multicentre validation trial in breast cancer, published by Krag and colleagues.\(^6\) The number of sentinel node biopsies per surgeon ranged from 16 to 51 with a total number of 443 procedures. The success rate differed significantly among the participating surgeons with a minimum of 79% and a maximum of 98% identification. The false-negative rate ranged from 0% to 29%. In a study of Cody et al., the identification rates of the four most experienced surgeons were between 92% and 96% and the false-negative rates between 7% and 20%.\(^5\) Bass and colleagues found that five surgeons obtained a 90% success rate (+/- 4.5%) after 23 sentinel node biopsies on average, and 95% (+/- 2.3%) after 53 sentinel node biopsies in breast cancer.\(^4\) In melanoma, the existence of a learning curve for sentinel node biopsy was demonstrated at The Netherlands Cancer Institute.\(^7\) The identification rate with the use of patent blue dye alone increased from 87% to 96% for one surgeon and from 76% to 86% for another surgeon. Almost all sentinel nodes (99.5%) were identified by adding the gamma-ray detection probe.

An institutional experience in sentinel node biopsy is made up by the experiences of the individual surgeons. A starting surgeon can either learn from his experienced colleagues or has to introduce the new procedure starting from scratch. The surgeon may find it helpful to attend a teaching course first.\(^8\) There are also several books and CD-ROMs that may help in the preparation.\(^9\)\(^-\)\(^14\)

This article discusses the effect of the number of patients in the learning phase for sentinel node biopsy on the precision with which the identification rate and sensitivity can be assessed and hence on the probability to correctly or incorrectly accept the quality of the performance of a surgeon or institution.

**Methods**

The identification rate is defined as the percentage of all patients undergoing a sentinel node procedure in whom a sentinel is identified. In case of a tumour-positive lymphatic basin, the sentinel node can be true-positive or false-negative. The false-negative rate is defined as the number of false-negative procedures divided by the sum of the true positive and false-negative procedures:
The sensitivity is defined as the complement of the false-negative rate:

\[
\text{Sensitivity} = \frac{\text{true-positive}}{\text{true-positive + false-negative}} \times 100\%
\]

Calculationss are done for four learning phases with a total number of sentinel node procedures of 25, 50, 75 and 150. The numbers of tumour-positive procedures corresponding with the chosen group-sizes are based on the assumption that 40% of the breast cancer patients will have lymph node metastases.\(^5\) This results in 10, 20, 30 and 60 tumour-positive procedures for the different learning phases. The acceptable long-term non-identification rate as well as false-negativity is assumed to be 5%. In addition, results for long-term false-negativity of 10%, 15% and 20% are also calculated for the groups of 20 and 60 tumour-positive procedures.

Critical values are defined for non-identification rate and false-negativity, in such a way that the quality of the procedure is accepted if both observed non-identification rate and false-negativity are lower than their critical value. Using the binomial distribution, it is then possible to calculate the probability that an observed non-identification rate or false-negativity is lower than its critical value, given a certain long-term non-identification rate or false-negative rate for the surgeon or centre. Depending on the assumed long-term value, the calculated probability is the chance to correctly or incorrectly accept the quality of a surgeon or centre. The chance of not accepting the quality is defined as the probability that an observed result is equal or higher than its critical value. Critical values to be used are given in the results section.

**Results**

Figure 1 illustrates the impact of the number of operations on the probability of failure to identify a sentinel node in a certain percentage of patients, given a 5% long-term non-identification rate. A surgeon with the inherent ability to find 95% of the sentinel nodes (probability of non-identification of 5%) still has a 13% chance of an observed non-identification rate of at least 10% (critical value) if the number of patients in a learning phase is only 25. This chance is 10%, 3% and 1% if the same surgeon performs 50, 75 and 150 procedures respectively. These probabilities represent the sum of the bars that are visible on the 10% level and higher levels in figure 1.

The first two probabilities (13% and 10%) of incorrectly rejecting the quality of the identification rate can be reduced to 5% if a critical value of 16% (N=25) or 12% (N=50) is used. But then of course the probability of inadvertently accepting a high long-term non-identification rate will be increased.
Figure 1. Binomial distribution showing the percentages of not-identified sentinel nodes based on the probability of non-identification of 0.05. The numbers on the Z-axes represent the four groups with 25, 50, 75 and 150 sentinel node procedures. The probability corresponds with the height of a bar and represents the chance of finding a certain non-identification rate.

Figure 2. Binomial distribution showing the false-negative rate in 10, 20, 30 and 60 tumour-positive (T+) sentinel node procedures based on the probability of a false-negative sentinel node of 0.05.
The chance of incorrectly accepting an actual identification rate of 85% in the different group sizes is 25% (N=25), 11% (N=50, N=75) and 3% (N=150). The critical value should be adapted from 10% to 4% (N=25) or 8% (N=50, N=75) to reduce the first two probabilities (25% and 11%) to less than 5%. However, the use of these critical values will result in a chance of 36% (N=25), 10% (N=50) and 8% (N=75) to reject a 95% long-term identification rate.

Similar histograms can be made for the false-negative rate (figure 2). A striking difference in comparison to figure 1 is the more extended range of values with the same probability of 0.05. This phenomenon is caused by the fact that the false-negative rate is based on a subgroup of all patients who undergo sentinel node biopsy, namely those with tumour-positive lymph nodes. Consequently, the group size of a learning phase should be based on the desired certainty about the safety of the procedure in terms of a low false-negative rate. For a surgeon with an actual false-negative rate of 5%, the chance of finding a false-negative rate of at least 10% (critical value) is 40%, 26%, 19% and 8% if the learning phase contains 10, 20, 30 and 60 tumour-positive procedures respectively (figure 2). These probabilities of incorrectly rejecting the sensitivity of the sentinel node procedure can be reduced to 5% if a critical value of 30% (N=10), 20% (N=20), 17% (N=30) or 12% (N=60) is used. However, this approach will increase the probability to accept a long-term false-negative rate that is too high.

![Figure 3. Binomial distribution showing the false-negative rate in a group of 20 axilla-positive patients based on four different probabilities of a false-negative (FN) sentinel node, namely 0.05, 0.10, 0.15 and 0.20.](image_url)

In figures 3 and 4, the probability to observe certain false-negative rates is displayed based on four different long-term probabilities of a false-negative procedure. Figure 3 shows a substantial overlap between the binomial distributions based on twenty
tumour-positive procedures when the probability of finding a false-negative sentinel node is increased with steps of 5%. For example, the chance of incorrectly reaching a favourable false-negative rate of less than 10% in these twenty patients is 18% for a surgeon with a long-term probability of false-negative procedures of 15%. On the other hand, a skilful surgeon with an actual 95% sensitivity has a 26% chance of accidentally finding an unfavourable false-negative rate of 10% or more in a group of 20 patients with involved lymph nodes. When the number of patients with lymph node metastases is increased to 60, the binomial distribution histograms become more separated from each other, enabling a more reliable conclusion about the ultimate false-negative rate (figure 4). The previously calculated chances in 20 tumour-positive procedures are now reduced from 18% to 10% and from 26% to 8%.

![Figure 4](image)

Figure 4. Binomial distribution showing the false-negative rate in a group of 60 tumour-positive sentinel node procedures based on four different probabilities of a false-negative (FN) sentinel node, namely 0.05, 0.10, 0.15 and 0.20.

Table 1 shows that if the chance of wrongly accepting a favourable result has to be 5% at the most, the critical value should be adapted to the number of positive procedures. In twenty tumour-positive procedures the critical value of the false-negative rate has to be 5% to have a less than 5% chance of incorrectly accepting a long-term probability of false-negativity of 15% or higher. In 60 tumour-positive procedures, the critical value can be increased to 8% to meet this requirement. With only ten tumour-positive procedures, this goal cannot be reached (table 1). In melanoma, the percentage of occult lymph node metastases is about 20%. A learning phase should therefore be twice as long as in breast cancer to reach the same statistical goals.
Table 1. Probability of incorrectly accepting a false-negative rate of 15% or higher in a learning phase with 10, 20, 30 and 60 tumour-positive procedures.

<table>
<thead>
<tr>
<th>Total no. of SN procedures</th>
<th>No. of tumour-positive SN procedures</th>
<th>Critical values* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>75</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>150</td>
<td>60</td>
<td>0</td>
</tr>
</tbody>
</table>

SN=sentinel node; *-Critical values for false-negativity are defined in such a way that the quality of the sentinel node procedure is accepted if the observed false-negative rate is lower than its critical value. The probability (%) of accepting the quality is calculated for chosen group-size and critical value using the binomial distribution.

Discussion

A realistic aim for the false-negative rate should be 5% or at least below 10%. Because small deviations from these percentages are already considered to be unacceptable, the confidence interval has to be narrow. Assuming that the ultimate false-negative rate is 5%, the chance of still finding a false-negative rate of at least 10% decreased from 40% to 8% when the number of tumour-positive procedures increased from ten to 60 (figure 2). This explains why Krag and colleagues could not draw firm conclusions from the differences in false-negative rate, although statistically different identification rates were found. The small number of patients with tumour-positive sentinel nodes in their study (four to sixteen per surgeon) leads to huge binomial confidence intervals between 0% and more than 60%.

This paper is focussed on the surgeon’s learning phase, but clearly other factors also play a role. Several studies have pointed out that a combined technique of both patent blue dye mapping and the use of a gamma ray detection probe after administration of a technetium-labelled colloid is superior to a single agent technique. Tracer migration from the dense dermal lymphatic network in melanoma results in a clearer visualisation of the sentinel node in comparison to breast cancer and probably influences the length of the learning curve in these two malignancies. There is increasing evidence that preoperative lymphoscintigraphy is helpful. It is important that the surgeon reviews the images together with the nuclear medicine physician. A pathologist who scrutinises the lymph node using serial sectioning and immunohistochemical staining is also essential. So, lymphatic mapping requires a team effort. A surgeon, no matter how skilled, will never find 95% of the sentinel nodes with a 95% sensitivity if not backed by a competent nuclear medicine physician and a competent pathologist. Different teams may have different results. However, a standardised technique within an institution allows evaluation and comparison of the performance of individual surgeons.

Do identification rate and sensitivity go hand in hand? Theoretically, re-routing of lymph flow to a neo-sentinel node because of tumour blockage does not hamper identification of the neo-sentinel node that may still be tumour-free. In other words,
such a false-negative case will not impair the identification rate. Krag described a
surgeon with a 98% success rate and a 27% (4 of 14) false-negative rate and another
surgeon with 79% identification and a false-negative rate of only 7% (1 of 14).\footnote{6} From
this point of view, it is not sufficient to consider only identification rate as a parameter
in determining the length of a learning curve. The clinical impact of a false-negative
sentinel node procedure seems more pronounced than not finding the node, which
will be followed by an axillary lymph node dissection in breast cancer patients. The
false-negative rate should be the most important factor in discussions about the
number of procedures that have to be completed before reliable conclusions can be
drawn about the performance of a team.

The term learning curve suggests that most failures occur in the initial phase of the
new procedure. If this is true, our calculations on the required number of patients
should be considered optimistic. The binomial distribution assumes a constant
failure probability. Valid statistical inference on the basis of this model has to be
restricted to the subgroup of later patients in the learning curve for whom this can be
assumed. Other models for which this assumption is not necessary will always lead
to less precision than inference from a binomial distribution with the same number
of patients. Curves of sentinel node identification in breast cancer published by Bass
et al. indeed show a high initial non-identification rate of more than 30% after a few
procedures.\footnote{4} Subsequently, the percentage of non-identification diminished
gradually and stabilised between 0% and 10% after about fifteen to 40 procedures. In
their recent experience with more participating surgeons, these investigators found
two other types of curves.\footnote{21} The first one was a late onset of unsuccessful sentinel
node biopsies. The second one was an increasing non-identification rate from the
beginning. Late failures were significantly associated with a low individual surgical
volume index (number of procedures per month per surgeon). These observations
support the idea that good results in the initial series of procedures are not always
indicative for future procedures.

In the experience of Cody and associates, false-negative cases occur early in the
experience of a surgeon.\footnote{5} Two of the four most experienced surgeons had their false-
negative biopsy in the first ten cases (case no. 1 and 6), and the other two in the next
ten cases (case no. 12 and 20). With exclusion of the first six biopsies of each surgeon,
the false-negative rate decreased from 11% to 5%. These observations are in contrast
to our own findings in 82 breast cancer patients who underwent sentinel node
biopsy with confirmatory axillary lymph node dissection, performed by three
surgeons. The sentinel node was false-negative in the 62\textsuperscript{nd} and 78\textsuperscript{th} procedure. It is
hard to draw conclusions from these findings, because of the low incidence and the
complex aetiology of false-negative procedures. The contribution of the surgeon to
the total false-negative rate of a multidisciplinary team in variable patient
populations based on different selection criteria cannot be determined exactly. One
aspect of the relation between false-negative results and the number of procedures is
indisputable: the false-negative rate will never be zero percent.

It is unclear how rigorously one should adhere to a certain false-negative rate. This
rate depends on how it is determined. Do false-negative biopsies have to declare
themselves during follow-up as is customary for a learning phase in melanoma patients? If completion regional lymph node dissection is done as is customary for a learning phase in breast cancer, how rigorously are metastases in non-sentinel nodes pursued? Is a truly complete nodal clearance carried out? Does the pathologist find all the nodes in the chunk of fatty tissue that the surgeon submits? A team with a skilful surgeon performing radical node dissections and a dedicated pathologist examining all non-sentinel nodes with serial sectioning and immunohistochemistry may find more tumour-positive non-sentinel nodes. Could it be that a better team will have a higher false-negative rate?

How many procedures should a surgeon perform in order to be qualified to implement lymphatic mapping in routine patient management? A total number of 20 to 30 procedures as proposed by Giuliano seems to be inadequate to reliably calculate the safety of the procedure in the hands of a specific surgeon and the other members of the multidisciplinary team. Morton suggested that an acceptable level of technical skill requires success in 30 to 50 melanoma patients and 60 to 80 cases of breast cancer. Our results show that at least 150 procedures, containing 60 with lymph node involvement, are needed to make a somewhat reliable conclusion about the quality of sentinel node biopsy. Even with this high number of procedures, there is a chance of inadvertently characterising the quality as sufficient or insufficient. It will take 750 patients with 300 tumour-positive basins to establish with 95% certainty that a surgeon who has a non-identification rate of 5% and a false-negative rate of 5% indeed has these capabilities within a range of 0% to 7%. It is unrealistic to pursue these standards. Increasing numbers of well-informed breast cancer patients nowadays are not willing to undergo a confirmatory axillary lymph node dissection when a nearby hospital has abandoned confirmatory axillary node dissection. These practical problems in controlled implementation of sentinel node biopsy were illustrated by a survey among surgeons in the Netherlands. Routine complete lymph node dissection in breast cancer patients was omitted by 43% of surgeons who had performed less than ten procedures, by 44% of surgeons with less than 25 and by 70% after a learning phase of less than 50 procedures. A survey among American surgeons who had attended a university-sponsored course showed that almost all surgeons (95%) did not complete 30 validation cases and 55% of them completed less than ten. In melanoma, the situation is even more complicated because the incidence of the disease is less, the percentage of involved basins is less, and confirmatory lymph node dissection is difficult to justify. Long-term follow-up is needed to determine the actual false-negative rate. Still, increasing numbers of melanoma patients wish to be informed about their lymph node status. An important aspect in the decision about the length of a learning curve is to determine a desired certainty that a result will be below a critical value. The chosen critical value has an important impact on that certainty particularly in relatively short learning phases as shown in this study. When the chance of incorrectly accepting an actual false-negative rate of 15% or higher has to be reduced to a minimum of 5%, the critical value has to be lower in smaller groups of patients (table 1); 5% in 20 tumour-positive procedures permitting no false-negative procedures.
Series from high-volume centres and experienced investigators confirmed that the sentinel node is the first lymph node reached by tumour cells in about 95% of all metastasised tumours with a small confidence interval. These series however are no guarantee of success for starting surgeons. Statistical analysis shows that a short learning phase may not be representative for the ultimate sensitivity and may conceal an unacceptable high recurrence rate in the future.

**Conclusions**

Identification rate and false-negative rate are the two parameters in the discussion about the learning phase. The false-negative rate is the most important one because of its clinical impact and the fact that it is based on a subgroup of patients which makes it the limiting factor. In breast cancer, the binomial distribution of the two variables for different group sizes makes clear that an unattainable number of procedures are required for a reliable judgement of the actual skills of a team performing sentinel node biopsy. The often-suggested number of 30 procedures is insufficient from a statistical point of view. At least 150 procedures, containing 60 with lymph node involvement, are needed to make a somewhat reliable conclusion about the quality of sentinel node biopsy. In short learning phases, the critical value of the identification and false-negative rate should be adjusted to reduce the chance of incorrectly accepting a low quality of performance of lymphatic mapping. The tumour type, technical aspects and the performance of other team members also determine a surgeon’s learning curve. It is impossible to make recommendations about the duration of the learning phase as a statistically sound learning phase is not attainable in clinical practice and a clinically practical learning phase lacks the desired statistical significance.

**References**


