Optimizing breast cancer survival models based on conventional biomarkers and stromal parameters

Dekker, T.J.A.; Dekker T.J.A.

7

Interobserver study among pathologists regarding the quantitative evaluation of lymph vascular space invasion (LVSI)

Manuscript in preparation

T.J.A. Dekker, E. Groen, J.R. Kroep, W.E. Mesker, R.A.E.M. Tollenaar,
M.J. van de Vijver, J. Wesseling, V.T.H.B.M. Smit

Introduction

The prognostic implication of lymph vascular space invasion (LVSI) has been investigated in multiple studies. This parameter has traditionally been investigated by stratifying tumors according to LVSI-negativity or -positivity [1-16]. However, published results by our group [17] and others [18] have shown that some form of quantitative evaluation of peritumoral LVSI might provide stronger clinical information for node-negative breast cancer patients. Determining LVSI in order to guide decisions regarding adjuvant therapies was initially recommended by the St. Gallen guidelines [19]. Although other guidelines still endorse determining LVSI for selecting breast cancer patients for adjuvant therapies [20], the most recent St. Gallen guidelines stated that vascular space invasion is not recommended for this purpose [21]. Although research has shown that quantitative assessment of this parameter can be valuable for risk prognostication in some otherwise low-risk breast cancer patients [17], quantitative LVSI assessment can only be implemented if reproducible scoring can be assured.

Little is known regarding the reproducibility and reliability of LVSI evaluation by pathologists. Published prognostic studies have not performed such analyses and interobserver statistics are rarely reported [2-4, 22]. Gilchrist et al. conducted a study investigating the interobserver variation among 9 pathologists and concluded that the variation between observers is too high to justify routine management of this parameter in breast cancer [23]. Secondly, no form of quantification was applied in this study. Small foci of LVSI might be more easily missed, and the concordance between observers might therefore increase if only tumors with 'extensive' LVSI are considered positive. To address these issues and their effect on the interobserver concordance, we have set up a study investigating the concordance rate between 4 dedicated breast pathologists (EG, MJvV, JW, VTHBMS) regarding extratumoral, quantitative LVSI assessment. For this purpose, 60 slides were observed by these four observers who all scored and semi-quantified LVSI as defined in a previous publication [17].

Methods

Study population

Standard H&E slides were retrieved for 60 tumors, corresponding patient and tumor characteristics are shown in table 1. The slides were selected to contain a mixture of tumors containing extensive LVSI, limited LVSI, and no LVSI. Cases were derived

from a cohort of the perioperative chemotherapy (POP) trial (EORTC 10854), which randomized 2795 patients with T1-T3, N0-2, and M0 breast carcinomas to receive one course of perioperative chemotherapy or no additional treatment. Previous analyses from our group regarding LVSI assessment described 441 premenopausal node-negative patients from this trial [17]. All tumors were negative for lymph node metastases, and were treated with either breast conserving surgery or radical mastectomy [24]. Estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), Ki-67 and p53 status were determined as previously described [25].

Scoring

After selection of cases, all slides were sent to four dedicated breast pathologists (EG, MJvdV, JW, VTHBMS). All observers were blinded to both scoring results of other observers and clinical outcome. Observers were asked to score LVSI lesions according to strict criteria. A lesion was considered as LVSI when it adhered to two of three the following criteria: 1) discordance between the contour of the suspected tumor embolus and lymph vessels, 2) presence of a clear endothelial lining of the lymph vessel, 3) a blood vessel in the direct vicinity of the suspected LVSI lesion (figure 1A). These criteria were based on previous publications and expert opinion on this matter [12, 23, 26-28]. If extensive retraction artefacts or tissue damage occurred in the surroundings, this was considered as sufficient reason not to classify a lesion as LVSI. Observers were asked to count the number of convincing LVSI lesions. Lesions that were found within one 10x magnification field were considered as one focus. Quantification was performed by multiplying the number of LVSI foci with the number of tumor cells within the *largest* tumor emboli, resulting in the LVSI tumor burden score (LVSI-TB). At scores below 60, the tumor was classified as LVSI positive, but LVSI-TB low. At scores equal to or exceeding 60, the tumor was classified as LVSI positive and LVSI-TB high.

The observers were also asked to mark all individual lesions on the microscopic slides that were considered to adhere to the LVSI criteria alongside their scores. The LVSI foci identified by individual observers were noted. Afterwards, all lesions that showed discordance between at least one observer were digitally photographed, and were again shown to all observers again in order to determine intra- and interobserver reliability for each individual suspected LVSI lesion. All observers were asked if this photographed lesions was considered as a convincing LVSI lesion in their opinion, and if not, what the reason was not to classify this lesion as LVSI.

Data analysis

Based on the scores by the individual observers, concordance and Cohen's kappa coefficient values for each individual LVSI-suspected lesion were calculated. In order to establish the consensus score, consensus for each individual lesion was assessed. When 3 or more observers considered a lesion to be LVSI, this was considered as an LVSI focus. The number of LVSI foci was then multiplied by the number of tumor cells in the largest tumor embolus (assessed by TJAD) to calculate the consensus score for each individual tumor (figure 2).

The topographical pattern for all suspected LVSI foci was assessed by one observer (TJAD) and stratified into one of three categories that were previously described by Orbo et al. [27] or into the 'other' category when no evident pattern was found (figure 3). Interobserver discordance for all individual lesion per topographical pattern was then assessed in order to determine whether these patterns affected interobserver variability. All statistical analyses were performed using IBM SPSS statistics version 20. Survival curves were calculated by using Kaplan-Meier statistics and compared via log-rank tests.

Results

Qualitative LVSI evaluation

We first sought to determine the inter-observer reproducibility of a qualitative evaluation of LVSI (LVSI-positive vs LVSI-negative). At least one LVSI suspected lesion was found by at least one observer in the peritumoral field of 33 tumors (55%). In total, the observers identified 92 different lesions that were suspected to be LVSI across these 33 tumors. A mean number of 58 individual LVSI foci were identified per observer (range 44 to 75). Only 27 out of the 92 lesions (29%) were identified by all observers, for all other suspected lesions at least one observer was discordant. The reasons for the 65 discordances could either by that the LVSI-suspected lesion was not seen by the observer, or that the observer had detected the lesion, yet decided this lesion was not convincing enough to be considered as an LVSI focus. Pictures of these 65 lesions were again presented to the observers to determine the inter-observer variation for these cases. In only 25 cases (38%) concordance between all 4 observers was reached when these cases were again presented to all observers. For all other cases, at least one observer remained discordant with the others (example shown in figure 1B). We examined the influence of the topographical patterns previously described by Orbo

et al. on the concordance between observers. All lesions were classified as one of four topographical categories, namely paravascular, interlobular, multiple lymphatic and other. Highest concordance was seen for the paravascular and multiple lymphatic patterns, and this was lower for the interlobular pattern, although this was not statistically significant (*P*=0.125). Of the 52 LVSI-suspected lesions were concordance was found after both scoring sessions, of which 38 were considered concordantly positive. The distribution of concordantly positive among the four topographical patterns was determined. Again, the interlobular pattern had the lowest amount of concordantly positive lesions (*P*=0.015).

These data infer that there is considerable variation concerning the assessment of individual LVSI foci. However, this does not necessarily imply that determining LVSI positivity cannot be reproducibly scored by observers. Considering the fact that LVSI foci commonly occur in multiple fields in the peritumoral area, as long as one of these foci is detected this would still lead to a reproducible score as LVSI-positive, even if not based on the same LVSI focus. We examined the concordance rate between the 4 observers concerning LVSI positivity. The Cohen's kappa coefficient values between observers ranged from 0.59 to 0.80 and the concordance between two observers ranged from 80% to 90% (table 2). Of all 60 tumors, 17 (28%) had at least one discordant result between all observers. In 6 cases, one observer had a LVSI-positive observation, which was not found by the 3 other observers. The opposite was true for 8 cases. The remaining tumor was scored as LVSI-positive by two observers, and negative by two.

Quantitative LVSI evaluation

A previous report from our group has shown that the LVSI-TB low and LVSI negative tumors are prognostically similar. As such, the most important clinical distinction is between the LVSI-TB high group on one hand and the LVSI-TB low/LVSI negative tumors on the other. We examined the reproducibility of this quantitative examination for all 4 observers. For 46 of 60 tumors (77%), concordant results were found for all 4 observers. For the discordant cases, 8 tumors were found to be LVSI-TB high by one observer while the other observers scored these tumors as either LVSI-TB low or LVSI negative. The opposite was true for 3 others tumors. Another 3 tumors were deemed LVSI-TB high by 2 observers, but not other 2 observers. Cohen's kappa values between observers ranged from 0.52 to 0.86 and concordance values ranged from 80% to 95% (table 3).

Based on the information provided by the observers concerning the individual LVSI images, a consensus was determined regarding the quantitative LVSI evaluation

for each individual tumor (figure 2). Although this study was not designed to study prognostic value of this parameter, the data show a decreased overall survival for all tumors that were determined as LVSI-TB high (*N*=15, 25%) by the consensus score (figure 3, *P*=0.050). Sensitivity and specificity values for identifying these consensus LVSI-TB high tumors, were calculated for each individual observer. Overall sensitivity and specificity for all observers combined were 83.3% (range 66.7-93.3) and 92.8% (range 80.0-97.8), with an overall concordance rate of 90.4% (81.7-96.7). Overall survival curves per the scoring of individual observers are shown in figure 4.

Scoring algorithm

Due to the relatively low concordance between individual observers, we investigated an algorithm for reliable LVSI scoring in clinical practice, implementing multiple observers to maximize detection of LVSI-TB high tumors. Although not all consensus LVSI-TB high tumors were identified as such by all observers in the primary scoring round, in all but one case was the tumor at least identified as LVSI positive. In the sole remaining case, the observer did not identify any LVSI lesions in the initial screening and this tumor was considered LVSI-TB high in the consensus score. We investigated the concordance between two observers regarding the identification of LVSI-TB high tumors. For this, we assessed all scoring results for all observer combinations for tumors where at least one observer detected a single LVSI lesion (*N*=187). When both observers agreed to LVSI-TB high status for any given tumor, this was 94.4% concordant with the consensus score. If both observers agreed to LVSI-TB low status for any given tumor, only 11.4% was considered a LVSI-TB high tumor by consensus. When there was discordance between both observers, 31.1% of those tumors were LVSI-TB high by consensus (table 4).

Discussion

In this study, we assessed the interobserver variation in both the quantitative and qualitative scoring of LVSI in the peritumoral tissues of breast cancer. Previous studies have shown that this parameter provides prognostic information for these patients. However, studies that have assessed the reliability of scoring this marker in breast cancer showed limited reproducibility [23, 27].

Quantitative evaluation of LVSI was first described by Colleoni et al. [18]. In their study, extensive LVSI was defined as “one or more foci of peritumoral vascular invasion in more than one tumor block”. In multivariate analyses, only tumors with extensive LVSI were significantly associated with decreased overall survival (compared

to tumors that contained no peritumoral LVSI, hazard ratio 3.55, 95% CI 1.24-10.1). In contrast, Mohammed et al. employed D2-40 staining to detect both intratumoral and peritumoral in a subsequent study, and concluded that no significant differences existed between different measures of extent regarding LVSI [11]. A subsequent study by our group utilized a different method for semi-quantification and showed increasingly poor prognosis for increasing amounts of LVSI, while tumors that contained little LVSI showed no significant difference compared to LVSI negative tumors [17].

No data regarding the reproducibility of quantitative LVSI scoring has been published to our knowledge. For this study, four dedicated breast pathologists assessed the presence of- and performed quantification of LVSI in the peritumoral area according to a previously published method [17]. In order to determine the reproducibility of this scoring method, Cohen's kappa interobsever values were calculated. Additionally, we compared individual scoring results with that of a consensus score. This consensus score was established by showing captured images of LVSI-suspected lesions to all observers and using consensus among these observers to identify LVSI-TB high tumors. Sensitivity and specificity values for individual observers were calculated by using this consensus scores as the gold standard. Other potential gold standards include the use of additional IHC stains such as CD31 or D2-40/podoplanin. However, these are not necessarily specific for lymph vessels [29, 30], but might be of use to resolve interobserver discordance. Overall sensitivity and specificity for all observers for identifying these LVSI-TB high tumors was a combined 83.3% (range 66.7-93.3) and 92.8% (range 80.0-97.8%), with an overall concordance rate of 90.4% (81.7-96.7%) with the consensus score. Cohen's kappa values *between* observers varied from 0.52 and 0.86, with concordance between 4 observers ranging from 80 to 95%.

We assessed whether this scoring performance might be improved by implementing a second observer. Since all consensus LVSI-TB high tumors were scored as LVSI positive by all observers except in one case, this seems a suitable initial screening for these tumors. After this initial screening by a single observer, implementation of a second observation might increase reliability. If both observers agree to LVSI-TB status, this agrees with a high consensus score in 94.4% of all hypothetical observer pairings in this study. In cases that both observers agree to a LVSI-TB low status, this is also low by consensus in 88.6% in all hypothetical pairings. However, in the cases of disagreement between both observers the percentage of consensus LVSI-TB high tumors was as high as 36.17%. It seems that additional diagnostics are appropriate in this group to aid the resolution of the interobserver variability. For this purpose, CD31 and/or podoplanin might be suitable [1].

The question remains for which patients quantitative LVSI assessment is warranted. In the case of patients that are already considered as unfavourable N0, it is unsure whether this parameter has any additional benefit. A previous study from our group showed that in unfavourable N0 patients, LVSI-TB high tumors are associated with an especially unfavourable prognosis. However, what additional treatment might be indicated for these patients is unclear. In the case of otherwise favourable N0 patients, the presence of extensive LVSI results in a prognosis that is similar to unfavourable N0 patients and it therefore seems valid to treat these patients as such [31]. We therefore recommend that the extent of LVSI should be determined in all otherwise favourable N0 patients via the scoring algorithm depicted in figure 5. Initial screening can be performed via one observer, and the aid of a second observer would only be needed when LVSI is detected by the first observer.

Limitations of our study include a relatively low number of cases. However, the selected series was enriched for cases with LVSI in order to enable the evaluation of the concordance for such lesions between the participating observers. Due to this relatively low number of included cases which were all from a previously published cohort, investigation of the previously reported association of quantitative LVSI quantification with patient survival was not the aim of this current study. Multivariate analyses regarding survival analyses were therefore not performed. Kaplan-Meier curves were however calculated to investigate the prognostic value of the consensus score and the LVSI scores of individual observers. We have shown the reproducibility of the assessment of this parameter among participating breast pathologists and also propose an algorithm for implementation in clinical practice. As part of this algorithm, we propose investigation of H&E slides and apply immunohistochemical staining in cases of discordance between both observers. Due to the unavailability of the tissue blocks for the tumors included in this study, this was not performed for this current investigation. To our knowledge, no data exists regarding the interobserver reproducibility of the use of D2-40- and CD31-stained slides for detecting LVSI presence and it is also unknown whether discordance between observers might be best resolved via discussion or whether IHC is necessary in these cases.

Data from our study shows that there is significant discordance between pathologists regarding individual LVSI foci. In this study we show that that sensitivity and specificity for quantitative LVSI scoring was 83.3% and 92.8% on average, with an overall concordance rate of 90.4%. Implementing a scoring system with a second observer might be used to increase scoring sensitivity although additional immunohistochemical tests seem indicated in some scenarios.

References

1. Braun M et al.: Detection of lymphovascular invasion in early breast cancer by D2-40 (podoplanin): a clinically useful predictor for axillary lymph node metastases. Breast Cancer Res Treat 2008, 112:503-511
2. Arnaout-Alkarain A et al.: Significance of lymph vessel invasion identified by the endothelial lymphatic marker D2-40 in node negative breast cancer. Mod Pathol 2007, 20:183-191
3. Britto AV et al.: Immunostaining with D2-40 improves evaluation of lymphovascular invasion, but may not predict sentinel lymph node status in early breast cancer. BMC Cancer 2009, 9:109
4. Ejlertsen B et al.: Population-based study of peritumoral lymphovascular invasion and outcome among patients with operable breast cancer. J Natl Cancer Inst 2009, 101:729-735
5. Hasebe T et al.: Histological characteristics of tumor in vessels and lymph nodes are significant predictors of progression of invasive ductal carcinoma of the breast: a prospective study. Hum Pathol 2004, 35:298-308
6. Ito M et al.: Significance of pathological evaluation for lymphatic vessel invasion in invasive breast cancer. Breast Cancer 2007, 14:381-387
7. Kato T et al.: The combination of angiogenesis and blood vessel invasion as a prognostic indicator in primary breast cancer. Br J Cancer 2003, 88:1900-1908
8. Lauria R et al.: The prognostic value of lymphatic and blood vessel invasion in operable breast cancer. Cancer 1995, 76:1772-1778
9. Lee AH et al.: Prognostic value of lymphovascular invasion in women with lymph node negative invasive breast carcinoma. Eur J Cancer 2006, 42:357-362
10. Marinho VF et al.: Lymph vascular invasion in invasive mammary carcinomas identified by the endothelial lymphatic marker D2-40 is associated with other indicators of poor prognosis. BMC Cancer 2008, 8:64
11. Mohammed RA et al.: Objective assessment of lymphatic and blood vascular invasion in lymph node-negative breast carcinoma: findings from a large case series with long-term follow-up. J Pathol 2011, 223:358-365
12. Pinder SE et al.: Pathological prognostic factors in breast cancer. III. Vascular invasion: relationship with recurrence and survival in a large study with long-term follow-up. Histopathology 1994, 24:41-47
13. Sabatier R et al.: Peritumoural vascular invasion: a major determinant of triple-negative breast cancer outcome. Eur J Cancer 2011, 47:1537-1545
14. Schoppmann SF et al.: Prognostic value of lymphangiogenesis and lymphovascular invasion in invasive breast cancer. Ann Surg 2004, 240:306-312

15. Viale G, G et al: Adverse prognostic value of peritumoral vascular invasion: is it abrogated by adequate endocrine adjuvant therapy? Results from two International Breast Cancer Study Group randomized trials of chemoendocrine adjuvant therapy for early breast cancer. Ann Oncol 2010, 21:245-254
16. Yamauchi C et al.: Accurate assessment of lymph vessel tumor emboli in invasive ductal carcinoma of the breast according to tumor areas, and their prognostic significance. Hum Pathol 2007, 38:247-259
17. Dekker TJ et al.: Quantitative assessment of lymph vascular space invasion (LVSI) provides important prognostic information in node-negative breast cancer. Ann Oncol 2013, 24:2994-2998
18. Colleoni M et al.: Prognostic role of the extent of peritumoral vascular invasion in operable breast cancer. Ann Oncol 2007, 18:1632-1640
19. Goldhirsch A et al.: Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. Ann Oncol 2007, 18:1133-1144
20. Senkus E et al.: Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013, 24 Suppl 6:vi7-23
21. Goldhirsch A et al.: Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol 2013, 24:2206-2223
22. Colleoni M et al.: Prognostic role of the extent of peritumoral vascular invasion in operable breast cancer. Ann Oncol 2007, 18:1632-1640
23. Gilchrist KW et al.: Interobserver variation in the identification of breast carcinoma in intramammary lymphatics. Hum Pathol 1982, 13:170-172
24. Clahsen PC et al.: Improved local control and disease-free survival after perioperative chemotherapy for early-stage breast cancer. A European Organization for Research and Treatment of Cancer Breast Cancer Cooperative Group Study. J Clin Oncol 1996, 14:745-753
25. Clahsen PC et al.: p53 protein accumulation and response to adjuvant chemotherapy in premenopausal women with node-negative early breast cancer. J Clin Oncol 1998, 16:470-479
26. Rosen PP: Tumor emboli in intramammary lymphatics in breast carcinoma: pathologic criteria for diagnosis and clinical significance. Pathol Annu 1983, 18 Pt 2:215-232
27. Orbo A et al.: Topographic criteria in the diagnosis of tumor emboli in intramammary lymphatics. Cancer 1990, 66:972-977
28. de M, I, B et al: Obvious peritumoral emboli: an elusive prognostic factor reappraised. Multivariate analysis of 1320 node-negative breast cancers. Eur J Cancer 1998, 34:58-65
29. Ren S et al.: D2-40: an additional marker for myoepithelial cells of breast and the precaution in interpreting tumor lymphovascular invasion. Int J Clin Exp Pathol 2011, 4:175-182

30. Yang Z et al.: Attenuated podoplanin staining in breast myoepithelial cells: a potential caveat in the diagnosis of lymphatic invasion. Appl Immunohistochem Mol Morphol 2009, 17:425-430
31. Goldhirsch A et al.: Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. Ann Oncol 2007, 18:1133-1144

Table 1. Characteristics of the tumors included in the study.

Parameter	*N*
Tumor size	22
pT1	29
pT2	1
pT3	8
Missing	
Tumor grade	20
I	14
II	21
III	5
Missing	
Hormone receptor positive	51
Hormone receptor negative	6
Missing	3
No HER2 overexpression	46
HER2 overexpression	12
Missing	2
Ki-67 low (< 20%)	11
Ki-67 high (> 20)	45
p53 low	48
p53 high	10
Deceased	8
Alive	50
Missing	2

Table 2. Cohen's kappa coefficient values for concordance between observers, regarding positive or negative for LVSI.

	Observer A	Observer B	Observer C	Observer D
Observer A	1			
Observer B	0.77 (88%)	1		
Observer C	0.80 (90%)	0.68 (85%)	1	
Observer D	0.70 (85%)	0.59 (80%)	0.69 (85%)	1

Table 3A. Concordance for individual lesions among observers stratified for topographical pattern (*P*=0.125).

Topographical pattern	Total number	Concordance	% Concordance
Paravascular	26	18	69.2
Interlobular	13	4	30.8
Multiple lymphatics	19	12	63.2
Other	34	18	52.9

Table 3B. Positive agreement for individual lesions among observers stratified for topographical pattern (*P*=0.015).

Topographical pattern	Total number	Positive agreement	% Positive agreement
Paravascular	26	16	65.1
Interlobular	13	4	30.8
Multiple lymphatics	19	10	52.6
Other	34	8	23.5

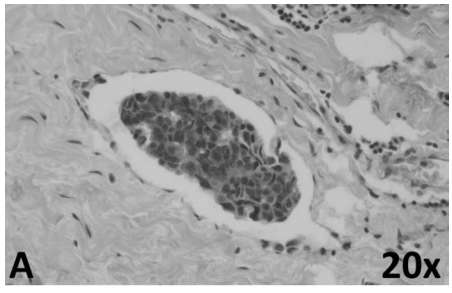
Table 4. Cohen's kappa coefficient values for concordance between observers, regarding quantitative evaluation for LVSI.

	Observer A	Observer B	Observer C	Observer D
Observer A	1			
Observer B	0.65 (85%)	1		
Observer C	0.76 (92%)	0.52 (80%)	1	
Observer D	0.86 (95%)	0.61 (83%)	0.71 (90%)	1

Table 4. All potential observer pairings for all tumors where at least one observer scored the tumor as LVSI positive were investigated. For all these scenarios, interobserver concordance regarding quantitative LVSI was investigated and compared to the consensus score. These data show that if two observers agree to an LVSI high score, this is generally concordant with the consensus score (in 94.4% of all instances). If both observers agree to an LVSI low score, this is generally concordant with the consensus score (88.6%).

Observer pairing	LVSI positive	Both LVSI-TB high/ consensus LVSI-TB high	One LVSI-TB high/ consensus LVSI-TB high	None LVSI-TB high/ consensus LVSI-TB high
A+B	32	14 / 13	9 / 1	9 / 1
A+C	33	13 / 13	3 / 1	17 / 1
A+D	30	11 / 10	5 / 4	14 / 1
B+C	33	11 / 10	12 / 3	10 / 2
B+D	28	10 / 10	6 / 3	12 / 2
C+D	31	13 / 12	10 / 2	8 / 1
Overall	187	72 / 68 (94.4%)	45 / 14 (31.1%)	70 / 8 (11.4%)

Figure 1. Two examples of suspected LVSI lesions.
1A. All observers agreed that this lesion adhered to all criteria for LVSI.
1B. For this lesion, only 2 out of 4 observers considered this as an LVSI lesion. The other two observers listed 'potential retraction artefact' and 'poor tissue quality' as the reason not to score this lesion as LVSI.



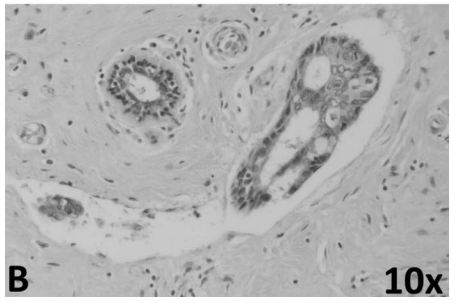


Figure 2. Determination of the LVSI-TB consensus score. All observers were shown captured images from suspected peritumoral LVSI lesions and were asked to determine whether those specific lesions were convincing LVSI foci. If at least 3 out of 4 observers agreed, the lesion was considered to be LVSI. For all LVSI lesions, the number of cells in the embolus were counted. The highest cell number was multiplied with the number of LVSI lesions in order to determine the LVSI-TB. If the outcome of this multiplication exceeded 60, the tumor was considered consensus LVSI-TB high. (A, B, C, D refer to individual observers).

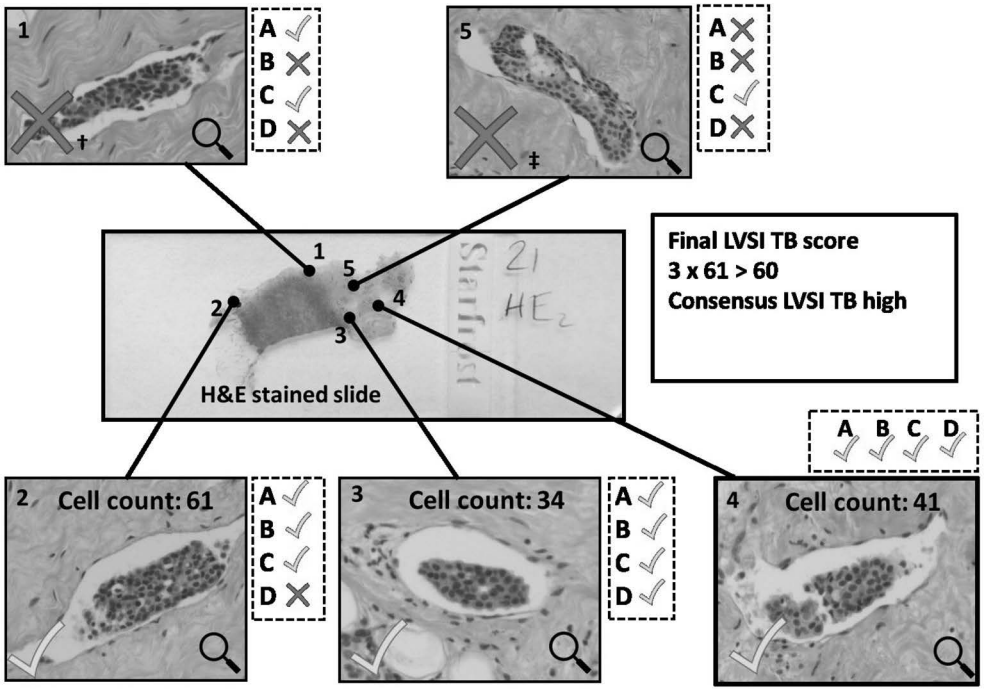


Figure 3A-C. Three topographical patterns of LVSI lesions, as previously described by Orbo et al. *3A*. The paravascular pattern: the LVSI suspected embolus is located in the direct vicinity of a arterial and venous blood vessel. Figure 3B. The interlobular pattern: the suspected tumor embolus is located within the direct vicinity of non-neoplastic lobules.
3B. The multiple lymphatics pattern: multiple lymphatic vessels with suspected tumor emboli.

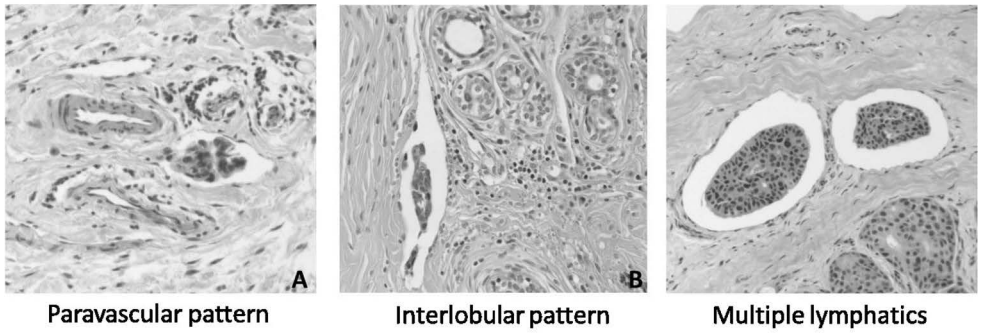


Figure 4. Overall survival plotted according to the LVSI-TB status via the consensus score. LVSI-TB high tumors (*N*=15) were associated with an overall survival time of 9.9 years (95% CI 7.907-11.952), compared to those with an LVSI-TB low status (*N*=45, overall survival 12.3 years (95% CI 11.7-13.0).

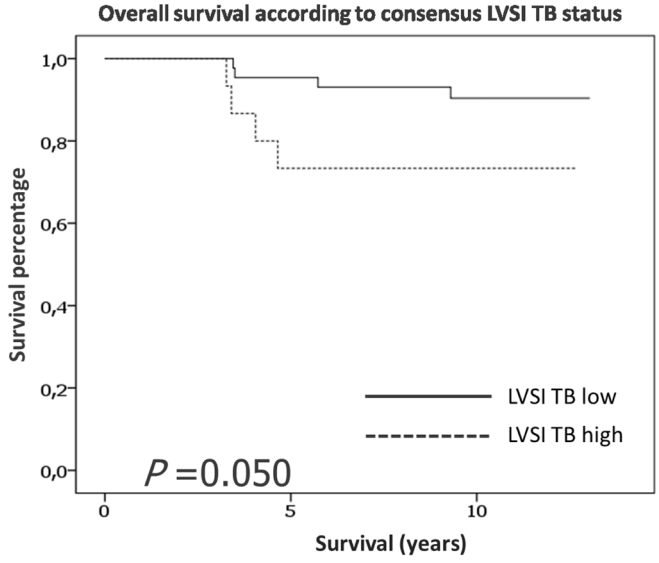


Figure 5. Overall survival according to LVSI-TB scoring according to the four observers.

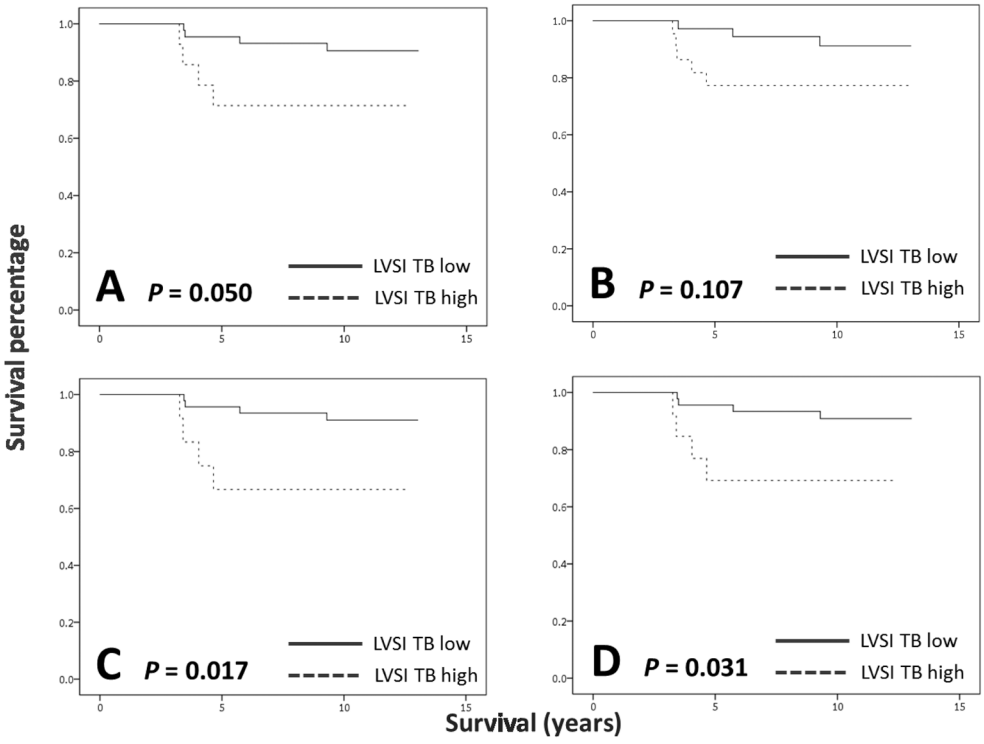


Figure 6. Suggested scoring algorithm to improve LVSI testing sensitivity by implementing a second observer.