Title page

Title of the paper: Mode of presentation and skin thickening on ultrasound may predict nodal burden in breast cancer patients with a positive axillary core biopsy

Running head title: Pre-operative nodal burden predictor

Author names: Wen Ling, WLC, Choong1, Andrew, AE, Evans2, Colin, CP, Purdie3, Huan, HW, Wang4, Peter, PTD, Donnan4, Brooke, BL, Lawson2, Jane, JM, Macaskill1

1. Ninewells Hospital and Medical School, Department of Breast Surgery, Level 6 Dundee, DD1 9SY, UK
2. Ninewells Hospital and Medical School, Department of Radiology, Level 6, Dundee, DD1 9SY, UK
3. Ninewells Hospital and Medical School, Department of Pathology Dundee, DD1 9SY, UK
4. Ninewells Hospital and Medical School, Population Health and Genomics, University of Dundee, DD1 9SY, UK

Corresponding author: Wen Ling Choong
Postal address: Ninewells Hospital and Medical School, Department of Breast Surgery, Level 6 Dundee, DD1 9SY, UK
Email address: choong7734@doctors.org.uk
Telephone number: 01382 633885

There were no sources of funding required for this publication. The manuscript is submitted as an original article. The paper is not based on any previous communication to a society or meeting.
Abstract

Objectives:
A number of pre-operative factors predicting nodal burden in women with breast cancer have recently been identified. The aim of this study is to assess if these factors independently influence nodal burden in women with a positive axillary core biopsy.

Methods:
All node positive patients detected on axillary core biopsy were identified in our cancer audit database. Mode of presentation, age, core tumour grade, core tumour type, ER and HER2 status were evaluated. Tumours were assessed for ultrasound (US) size, distance of tumour-to-skin, presence of invasion of skin and diffuse skin thickening. Axillary lymph nodes were assessed for cortical thickness and presence of US replaced nodes. Statistical significance was ascertained using univariate logistic regression. A predictive model was produced following a multiple logistic regression model incorporating cross validation and assessed using receiving operating characteristic (ROC) curve.

Results:
115 patients’ data were analysed. Patients referred because of symptoms (70% vs 38%, p=0.005), and those with US skin thickening (87% vs 59%, p=0.055) have higher nodal burden than those referred from screening or without skin thickening. These factors were significant after multivariate analysis. The final predictive model included mode of presentation, US tumour size, cortical thickness and presence of US skin thickening. The area under curve (AUC) is 0.77.

Conclusion:
We have shown that mode of presentation and US skin thickening are independent predictors of high nodal burden at surgery. A model has been developed to predict nodal burden pre-operatively, which may lead to avoidance of ANC in patients with lower nodal burden.

Advances in knowledge: Method of presentation and skin involvement/proximity to skin by the primary tumour are known to influence outcome and nodal involvement respectively but have not been studied with regard to nodal burden. We have shown that mode of presentation and skin thickening at US are independent predictors of high nodal burden at surgery.

(303 words)
Abstract

Objectives:
A number of pre-operative factors predicting nodal burden in women with breast cancer have recently been identified. The aim of this study is to assess if these factors independently influence nodal burden in women with a positive axillary core biopsy.

Methods:
All node positive patients detected on axillary core biopsy were identified in our cancer audit database. Mode of presentation, age, core tumour grade, core tumour type, ER and HER2 status were evaluated. Tumours were assessed for ultrasound (US) size, distance of tumour-to-skin, presence of invasion of skin and diffuse skin thickening. Axillary lymph nodes were assessed for cortical thickness and presence of US replaced nodes. Statistical significance was ascertained using univariate logistic regression. A predictive model was produced following a multiple logistic regression model incorporating cross validation and assessed using receiving operating characteristic (ROC) curve.

Results:
115 patients’ data were analysed. Patients referred because of symptoms (70% vs 38%, p=0.005), and those with US skin thickening (87% vs 59%, p=0.055) have higher nodal burden than those referred from screening or without skin thickening. These factors were significant after multivariate analysis. The final predictive model included mode of presentation, US tumour size, cortical thickness and presence of US skin thickening. The area under curve (AUC) is 0.77.

Conclusion:
We have shown that mode of presentation and skin thickening are independent predictors of high nodal burden at surgery. A model has been developed to predict nodal burden pre-operatively, which may lead to avoidance of ANC in patients with lower nodal burden.

Advances in knowledge: Method of presentation and skin involvement/proximity to skin by the primary tumour are known to influence outcome and nodal involvement respectively but have not been studied with regard to nodal burden. We have shown that mode of presentation and skin thickening at US are independent predictors of high nodal burden at surgery.

(303 words)
Introduction

Around 50% of patients with axillary node metastases from breast cancer have these diagnosed pre-operatively [1]. It is known that pre-operative node positive patients are found to possess higher nodal burden of disease compared to patients diagnosed surgically on sentinel lymph node biopsy (SLNB) [2, 3]. However, a significant number of patients with a pre-operative diagnosis of axillary metastases also have a low nodal burden [4].

Current standard practice is for patients with a positive pre-operative axillary core biopsy to proceed directly to axillary node clearance (ANC). As ANC is associated with high morbidity, there is an increasing move to steer away from ANC where there is no survival advantage to patients [5, 6].

Evidence from a landmark randomised trial (ACOSOG Z0011) has shown that not all node positive patients require ANC [7]. In women with a low burden of axillary disease i.e. one or two positive lymph nodes positive on SLNB, the AMAROS trial has shown that axillary radiotherapy provides comparable axillary control to ANC and with a lower morbidity [5, 8]. It would be advantageous therefore, to know preoperatively which patients with a diagnosis of axillary metastasis were likely to have a high or low burden of disease in order to guide planning of surgical procedures.

Two recent studies have shown that preoperative factors associated with high nodal burden with independent significance include the number of abnormal nodes seen on US and the node cortical morphology on US [9, 10]. Other factors which were significant on univariate analysis but not were not significant on multivariate analysis included needle biopsy (fine needle aspiration or core needle biopsy) was performed (p<0.0001), presence of effacement of fatty hilum (p<0.0001), larger breast tumour size measured on US (p<0.0001), reported abnormal LN maximum cortical thickness (p=0.0002), higher tumour grade (p=0.0001), presence of lymphovascular invasion (p=0.0038) and positive HER2 (p=0.0262) [9]; larger US short diameter (SD) of axillary lymph nodes (p=0.009), presence of MRI cortical morphological changes (p<0.001), larger MRI SD (p=0.025) and long diameter (LD) (p=0.036) of axillary lymph nodes [10]. The study by Kim et al had a number of limitations: (i) only 19 patients (6.1%) had a high nodal burden versus 293 patients (93.9%) having low
nodal burden, (ii) not all patients with positive axillae underwent ANC to determine the total number of positive nodes and (iii) only two-thirds (312 out of 451 patients) of patients had US meaning the results are liable to bias [10]. The low nodal burden group included patients who were node negative as well as those with 1-2 nodes positive. The study by Lim et al again included women who were node negative in the low burden group but the study did include a large number of women with a high nodal burden and all women with positive axillae appeared to have had ANC [9].

Recent studies have shown that skin involvement and a short distance from the tumour to the skin are associated with high rates of nodal involvement [11, 12]. These studies did not, however, look at the nodal burden in those patients with axillary metastases. A number of studies have shown that measuring breast skin thickness is accurate and reproducible [13, 14]. Screening detection is associated with a good outcome even when other pathological variables are taken into account. This is reflected in the widely used PREDICT score takes patient source of referral into account [15]. Neither of these two previous studies looked at presence of diffuse skin thickening or patient source at predicting nodal burden.

The aims of this study are, therefore, to assess pre-operative predictors of nodal burden specifically in those women who have positive axillary core biopsies (the first study to do so) and to see if US skin involvement and patient source of referral are associated with nodal burden.

Methods
A retrospective search of our local electronic cancer audit database for all node positive patients detected on axillary core biopsy between the 1st January 2014 and 31st December 2016 was carried out. Ethical approval was obtained. Exclusion criteria were patients who had on axillary core biopsy disease which on a node excision would be categorised as isolated tumour cells or micrometastatic disease. Patients who underwent neoadjuvant chemotherapy, neoadjuvant endocrine therapy or those who did not proceed to ANC were also excluded. This yielded a study group of 115 with a median age
of 62 (range 38-90) who proceeded to ANC after they were found to have nodal metastasis on core biopsy. This was standard local practice during the time frame of this study.

The preoperative factors assessed included mode of presentation (screen-detected versus symptomatic), patient’s age, tumour grade on core, tumour type on core, ER status and HER2 status. There was a retrospective review of ultrasound (US) images of the primary tumour and axillary lymph node. This was performed by an experienced radiologist who was blinded to the number of positive lymph nodes found on pathological assessment of the ANC specimen. The US images of the primary tumours were assessed for US size, distance of tumour to skin, presence of direct invasion of skin and presence of diffuse skin thickening over the tumour. The US images of the axillary lymph nodes were assessed for cortical thickness and the presence of replaced nodes (absence of the hilum). It was not possible to accurately estimate the number of abnormal nodes seen on US in this retrospective study.

The level of nodal burden at ANC was defined as low if 1-2 nodes were involved and heavy if 3 or more nodes were involved with macrometastases (>2mm). We considered the skin to be thickened if it was above 2.5mm at US. Figure 1 demonstrates measurement of skin thickness on US [16]. Figure 2 shows US images of radiological skin invasion of the tumour.

The association between the level of nodal burden and each of the candidate risk factors were initially assessed by performing univariate logistic regression before fitting the 12 candidate factors into the multiple logistic regression model with outcome as the level of nodal burden at ANC (defined as low if 1-2 nodes and heavy if 3 or more nodes). Statistical programme used to perform statistical analysis include using VassarStats (https://vassarstats.net) and R3.6.1.

**Results**

**Patient characteristics**

One hundred and fifteen patients proceeded to ANC after they were found to have nodal metastasis on core biopsy. All patients were female. The median age at diagnosis of breast cancer was 62 years (range
43 (37%) patients had 1-2 nodes positive (Group A) while 72 (63%) patients had 3 or more nodes positive (Group B). Patient characteristics in relation to nodal burden are presented in Table 2.

Mode of presentation

Patients referred through breast screening had a lower frequency of high nodal burden compared to women with symptoms (10/26 (38%) vs. 62/89 (70%) respectively, p=0.005).

Core biopsy

Patients with HER2 negative disease had a significantly lower frequency of high nodal burden (51/91 (56%) vs. 21/24 (81%), p= 0.032). Core tumour grade and ER status did not predict nodal burden. Core biopsy characteristics in relation to nodal burden are presented in Table 2.

Imaging factors

A subgroup of 50 patients have had their imaging reviewed by a second radiologist. The results are shown in Table 1. This confirms the excellent reproducibility of the metrics used. All US were carried out by 6 radiologists and 1 advanced practitioner of more than 5 years of experience of breast and axillary US performed the scan using the same model of US machine.

i. US images of the primary tumours

Smaller US tumour size (p=0.009) was significantly associated with a low frequency of high nodal burden. The presence of US skin thickening has a borderline association with high nodal burden. 13/15 (87%) vs 59/100 (59%) p=0.055. Direct skin invasion at US was not associated with heavier node positivity. Imaging factors in relation to nodal burden are presented in Table 3.

ii. US images of the axillary lymph nodes

Thinner LN cortical thickness (p= 0.002) was associated with a low frequency of high nodal burden. The percentage of patients with a high nodal burden was 33.3% for 2-3mm, 48.0% for 3.1-4mm and 68.2% for 4.1-5mm and 75.0% for >5mm. The presence of replaced nodes was not associated with a high frequency of high nodal burden. See Table 3.
Final predictive model

Table 4 gives the odds ratios for the factors in the final predictive model. Those factors maintaining independent significance were mode of presentation, US skin thickening, US tumour size and US node cortical thickness. The discriminatory power for the final predictive model was $c = 0.77$ (sensitivity = 0.51, specificity = 0.81). The ROC curve is shown in Figure 3.

Discussion

ANC has been utilised as the standard treatment for axillary metastasis in breast cancer for many years. It was not until recently that practices have changed in light of influential clinical trials such as AMAROS and ACOSOG Z0011. The AMAROS trial was the first study to suggest that ANC is not needed in all patients with positive nodes. At 6.1 median years’ follow up, there were no significant differences in the 5-year overall survival (93.3% vs 92.5%), disease-free survival (86.9% vs 82.7%) or loco-regional recurrences (0.43% vs 1.19%) between patients with T1-T2 breast cancer and no more than 2 positive SLNs, who were randomised to ANC versus axillary radiotherapy [8]. There was twice as much morbidity (i.e. lymphoedema) associated with ANC [8]. Authors of the ACOSOG Z0011 study, concluded that in patients with 1-2 nodes positive treated with breast conservation surgery, the use of sentinel node biopsy alone may be adequate and these patients may not require ANC [7]. In that randomised trial, female patients with T1-T2 breast cancer undergoing wide local excision and radiotherapy with no palpable lymphadenopathy and no more than 2 positive sentinel lymph nodes (SLNs) were randomised to ANC versus no further axillary surgery and at 9.3 median years there was no difference in overall survival (86.3% vs 83.6%; $p=0.02$) between the two groups. Similarly, there was no difference in disease-free survival (80.2% vs 78.2%; $p=0.32$) or regional recurrence (83.0% vs 81.2%; $p=0.41$) [17]. These trials, therefore, demonstrated that patients with limited axillary metastasis could avoid ANC without compromising overall survival and loco-regional outcomes.

In this study, we have found patients mode of presentation and US skin thickening are preoperative factors which are independently associated with nodal burden found on ANC in patients who had a
preoperative diagnosis of axillary metastasis on core biopsy, we also found US tumour size and US
cortical thickness to be independent predictors of nodal burden.

Two recent studies have shown that the number of US abnormal axillary nodes and US node
 morphology (including cortical thickness) are also independently associated with high nodal burden at
ANC. As cortical thickness is a continuous variable, the cut off value used to prompt biopsy can be
adjusted upwards to ensure a reduction in the pre-operative diagnosis of nodal metastases in women
with low nodal burden. As only one-third of patients has a cortical thickness of less than 3mm have a
high nodal burden, ANC clearance in this group may represent overtreatment. This data also supports
using a 3mm cut-off for biopsy for abnormal nodes. The two previous studies differed from our current
study as they included node negative women in the low nodal burden group, whilst our current study is
the first to look at predictors of nodal burden in women with positive axillary core biopsies only and is
the first study from a European Breast centre.

Regarding US skin thickening, all clinical inflammatory cancers were excluded from this study, and
these cases in this study did not have clinically apparent skin thickening. Previous studies have
identified that the distance from the tumour to the skin is related to nodal metastases but we did not find
this significant in our study [18]. The presence of skin thickening is postulated to be due to skin oedema
secondary to lymphatic obstruction by tumour cells. The high rate of nodal metastases is probably as a
result of tumour entering the florid dermal lymphatic and venous plexuses. Such a mechanism of
tumour dissemination has long been established in malignant melanoma [19, 20, 21].

More importantly, we have found that patient mode of presentation is an independent predictor of nodal
burden even when other factors such as US tumour size are taken into account. The reason for this is
unclear. This finding is however similar to the previous studies showing that screen detection is an
independent good prognostic factor even when multiple pathological factors are taken into account [22,
23].
Multiple nomograms and scoring systems have been developed to help predict the risk of metastases in non-sentinel lymph nodes in patients with a positive sentinel lymph node. These include the Memorial Sloan-Kettering Cancer Centre, MD Anderson Cancer Center, Stanford, Tenon and Saidi nomograms and scoring systems [24, 25]. The two recent studies of preoperative prediction of nodal burden in women with a positive axillary core biopsy found that number of abnormal nodes was a strong predictor. The number of abnormal nodes at axillary US could not be assessed in this retrospective analysis. Using the factors identified in our study a predictive model gives an AUC of 0.77. This is the same as one the studies which included the number of abnormal nodes on US. This suggests that a model including the number of abnormal nodes and the factors we have identified should give a model with an even better performance. A prospective, multi-centre study would be useful in confirming the validity of this approach before clinical practice can change.

**Conclusion**

The present study showed that US skin thickening, mode of referral, US tumour size and US cortical thickness are independently associated with nodal burden found on ANC in patients with a pre-operative diagnosis of axillary metastasis on core biopsy. When these factors added to the number of abnormal nodes on US, a model could be constructed to predict nodal burden pre-operatively, leading to an avoidance of ANC in patients with a lower nodal burden.

**Conflict of interest and financial disclosure statements**

The authors of this publication do not have any conflict of interest or financial relationships to disclose. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
References


**Figure**

**Fig. 1** US image demonstrates US skin thickening. The bold white arrow measures more than 2.5mm therefore classified as skin thickening.
**Figure**

**Fig. 2** US image shows direct skin invasion by the tumor. The white arrow points at skin invasion of the tumor.
Fig. 3 ROC curve for the final predictive model. The AUC is 0.77.
Tables

Table 1. USS results reproducibility

<table>
<thead>
<tr>
<th>USS parameters</th>
<th>Agreement/ICCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replaced node</td>
<td>0.80*</td>
</tr>
<tr>
<td>Skin thickening</td>
<td>0.86*</td>
</tr>
<tr>
<td>Skin involvement</td>
<td>0.96*</td>
</tr>
<tr>
<td>ICCC cortical thickness</td>
<td>0.94**</td>
</tr>
<tr>
<td>ICCC distance to skin</td>
<td>0.90**</td>
</tr>
<tr>
<td>ICCC tumour size</td>
<td>0.77**</td>
</tr>
</tbody>
</table>

*Agreement

**ICCC- intraclass correlation coefficient (values above 0.75 are classified as excellent)
Table 2. Patient Characteristics

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Group A (1-2 node(s)) (n = 43)</th>
<th>Group B (≥ 3 nodes) (n = 72)</th>
<th>Univariate OR (95% CI)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (median with IQR)</td>
<td>60 (49-69)</td>
<td>64 (54-74)</td>
<td>1.02 (0.99 – 1.06)</td>
<td>0.132</td>
</tr>
<tr>
<td>Mode of presentation**</td>
<td>Screening</td>
<td>16 (62%)</td>
<td>10 (38%)</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Symptom</td>
<td>27 (30%)</td>
<td>62 (70%)</td>
<td>3.67 (1.48 – 9.13)</td>
</tr>
<tr>
<td>Tumour grade</td>
<td>1 or 2</td>
<td>19 (40%)</td>
<td>28 (60%)</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>24 (35%)</td>
<td>44 (65%)</td>
<td>1.24 (0.58 – 2.68)</td>
</tr>
<tr>
<td>Tumour type</td>
<td>IDC</td>
<td>32 (34%)</td>
<td>63 (66%)</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Lobular</td>
<td>8 (47%)</td>
<td>9 (53%)</td>
<td>0.57 (0.20 – 1.62)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3 (100%)</td>
<td>0 (0%)</td>
<td>0.00 (0.00 - Inf)</td>
</tr>
<tr>
<td>ER status</td>
<td>Positive</td>
<td>35 (41%)</td>
<td>51 (59%)</td>
<td>0.56 (0.22 – 1.39)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>8 (28%)</td>
<td>21 (72%)</td>
<td>--</td>
</tr>
<tr>
<td>Her2 status</td>
<td>Positive</td>
<td>3 (15%)</td>
<td>17 (85%)</td>
<td>4.12 (1.13 – 15.02)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>40 (42%)</td>
<td>55 (58%)</td>
<td>--</td>
</tr>
</tbody>
</table>

* P values are obtained from t-test for age and from Chi-squared test for tumour grade, tumour type, ER.

** Included in the final predictive model
Table 3. Imaging factors.

<table>
<thead>
<tr>
<th>Imaging Factors</th>
<th>Group A (1-2 node(s))</th>
<th>Group B (≥ 3 nodes)</th>
<th>Univariate OR (95% CI)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>US tumour size (mm)</strong>**</td>
<td>Median (IQR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19 (14-25)</td>
<td>24 (17-35)</td>
<td>1.06 (1.01 – 1.10)</td>
<td>0.009</td>
</tr>
<tr>
<td><strong>US cortical thickness</strong>**</td>
<td>Median (IQR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.00 (3.45-5.40)</td>
<td>5.30 (4.10-8.15)</td>
<td>1.37 (1.13 – 1.67)</td>
<td>0.002</td>
</tr>
<tr>
<td>Missing (%)</td>
<td>0 (0%)</td>
<td>1 (1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distance from tumour to skin (mm)</strong></td>
<td>Median (IQR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.4 (3.1 – 8.0)</td>
<td>4.3 (1.5 – 6.9)</td>
<td>0.93 (0.83 – 1.03)</td>
<td>0.164</td>
</tr>
<tr>
<td><strong>Presence of direct invasion of skin</strong></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 (33%)</td>
<td>10 (67%)</td>
<td>1.23 (0.39 – 3.86)</td>
<td>0.728</td>
</tr>
<tr>
<td></td>
<td>38 (38%)</td>
<td>62 (62%)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Presence of US skin thickening</strong>**</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (13%)</td>
<td>13 (87%)</td>
<td>4.52 (0.97 – 21.09)</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>41 (41%)</td>
<td>59 (59%)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Replaced node</strong></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17 (29%)</td>
<td>42 (71%)</td>
<td>2.14 (0.99 – 4.63)</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>26 (46%)</td>
<td>30 (54%)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

* P values are obtained from Wilcoxon rank sum test for USS tumour size, USS cortical thickness, and distance from tumour to skin, and from Chi-squared test for presence of direct invasion of skin, presence of diffuse skin thickening, and replaced node.

** Included in the final predictive model
Table 4. Final predictive model for the level of nodal burden (high vs low).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient estimate</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of presentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptom</td>
<td>0.927</td>
<td>2.53 (0.88 – 7.25)</td>
</tr>
<tr>
<td>USS tumour size</td>
<td>0.067</td>
<td>1.07 (1.02 – 1.13)</td>
</tr>
<tr>
<td>USS cortical thickness</td>
<td>0.314</td>
<td>1.37 (1.09 – 1.72)</td>
</tr>
<tr>
<td>Presence of diffuse skin thickening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2.280</td>
<td>9.78 (1.01 – 94.62)</td>
</tr>
<tr>
<td>No*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Criterion for references