Abstract
Invasive breast cancer is the second most common cancer worldwide. It is known to metastasise to the regional axillary lymph nodes but there has been debate over what is the best way to stage and treat the axilla in patients presenting with primary breast cancer. Multiple trials over the last two decades have led to a change in practice from routine axillary lymph node dissection to sentinel lymph node biopsy in patients who are clinically lymph node negative preoperatively. This has resulted in new questions regarding subsequent treatment of some patients. This review will critically appraise the evidence on axillary treatment in patients with low burden axillary disease and highlight limitations of relevant randomised controlled trials.

1. Introduction
Breast cancer is the second most common cancer worldwide with an estimated 1.6 million new diagnoses worldwide in 2012.1 Traditional surgical treatment of invasive breast cancer involved axillary lymph node dissection (ALND), which served as both a regional staging procedure and regional treatment. The widespread adoption of sentinel lymph node biopsy (SLNB), led to debate about the ongoing need for ALND, particularly in patients with low burden axillary disease.

Routine SNLB has replaced routine axillary lymph node dissection in patients with clinically negative axillary nodes preoperatively, but this has left interesting questions as to how to treat patients afterwards. The results of the American College of Surgeons Oncology Group Z0011 trial in women who underwent breast conserving surgery (BCS) plus whole breast irradiation (WBI) suggested that axillary treatment was potentially over-treating patients.2 However questions have been raised as to the robustness of the methodology and results of the Z0011 trial. This article will review
the Z0011 trial against the background of the most influential papers containing other RCT evidence regarding axillary treatment in primary breast cancer. In doing so, it will address five questions:

1. Is there a need for ALND in SLN negative patients
2. Is there a need for ALND in patients with a SLN micrometastasis?
3. Is there a need for ALND in patients with SLN macrometastasis?
4. Is axillary radiotherapy equivalent to ALND in patients with a SLN macrometastasis?
5. In women with node positive or high risk node negative breast cancer who have undergone BCS + WBI (+ ALND if node positive), does the addition of regional nodal irradiation (RNI) to WBI improve outcome, where all patients were treated with adjuvant systemic therapy?

2. Question 1: Is there a need for Axillary lymph node dissection in Sentinel Lymph Node negative patients. (NSABP B32 trial, ClinicalTrials.gov NCT 00003830)

The theoretical risk of SLNB compared to ALND is that SLNB alone could miss regional lymph node metastases to the axilla and hence under stage the disease. This would result in ‘under-treatment’ of the axilla with negative impact on long term survival. NSABP B32 was a trial carried out to address these concerns. This large, multicentre randomized controlled trial recruited 5611 patients with invasive breast cancer and randomized them to two groups which were well matched for patient demographic and tumour biology criteria. Group 1 had underwent routine ALND after the sentinel node was sampled no matter the result, whilst Group 2 only underwent ALND if the sentinel node was positive.

Overall there were 3986 patients who were lymph node negative on SLNB. Of these, 2011 were in Group 1 and went on to have ALND and 1975 were in Group 2 who had no further axillary surgery. There was no difference in overall survival (the primary endpoint), disease free survival (DFS) or regional control between the two groups. Morbidity was significantly lower in the SNLB only group. It should be noted that while the Log-rank comparison of overall survival in Groups 1 and 2 showed no evidence of a difference (p=0.12), the analysis yielded an unadjusted hazard ratio of 1.20 (95% CI 0.96–1.50) – i.e. 20% difference in overall survival in favour of the ALND which was statistically non-significant. The overall survival for both groups was 91.8% (95% CI 90.4–93.3) in group 1 and 90.3% (88.8–91.8) giving an absolute difference in OS of 1.5%. It should also be noted that in this trial, the false negative rate following SLNB was 9.8% (95% CI 7.8–12.2) with an overall success rate of 97.2% (95% CI 96.4–97.7). Despite around 10% of patients in the SLNB only group being lymph node positive, the study met its non-inferiority endpoint. At the same time it could be argued that the non-significant increase of 20% all cause mortality in patients treated by SLNB alone may have been linked to ~10% of these patients having positive nodes (based on staging data from the ALND arm of the study).

After longer follow-up, the overall survival at 10 yrs, published in an abstract of a meeting presentation, reported that there continued to be no significant difference in OS between the two groups (HR: 1.11, p = 0.27). 10 yr Kaplan-Meier (K-M)
estimates for OS are 87.8% for SNR alone and 88.9% for SNR + AD – ie an absolute difference of 1.1%.

In summary, NSABP B32 provides good level 1 evidence, from a large, appropriately powered phase 3 RCT that SLNB is a reliable technique for staging the axilla and routine axillary clearance does not provide a survival benefit to SNLB node negative patients.

3. **Question 2: Is there a need for ALND in patients with a SLN micrometastasis? (IBCSG 23-01 trial, ClinicalTrials.gov NCT00072293)**

Since the adoption of SNLB, controversy has surrounded the management of patients who have a micrometastasis (≤2mm). IBCSG 23-01 randomised 931 patients with micrometastasis on SNLB to either undergo ALND or no local treatment to the axilla in a 1:1 ratio. This was a multicentre, randomised, non-blinded, non-inferiority, phase 3 trial, in which patients were eligible if they had clinically non-palpable axillary lymph node(s) and a primary tumour of 5 cm or less and who, after SLNB, had one or more micro-metastatic (≤2 mm) sentinel lymph nodes with no extracapsular extension. The primary endpoint was disease-free survival. Non-inferiority was defined as a hazard ratio (HR) of less than 1.25 for no axillary dissection versus axillary dissection. This means by the omission of ALND for patients with micrometastases that a survival difference of up to 25% more would be accepted as ‘confirming’ that SLNB was non-inferior to ALND. This is a very large margin to regard as an acceptable hazard ratio for a group of patients with only micrometastases who were otherwise receiving optimal anti-cancer care.

After the exclusion of three patients, 464 patients were in the axillary dissection group and 467 patients were in the no axillary dissection group. With a median follow-up of 5.0 (IQR 3.6–7.3) years, the study reported 69 disease-free survival events in the axillary dissection group and 55 events in the no axillary dissection group. Breast-cancer-related events were reported in 48 patients in the axillary dissection group and 47 in the no axillary dissection group (ten local recurrences in the axillary dissection group and eight in the no axillary dissection group; three and nine contralateral breast cancers; one and five regional recurrences; and 34 and 25 distant relapses). Other non-breast cancer events were recorded in 21 patients in the axillary dissection group and eight in the no axillary dissection group (20 and six second non-breast malignancies; and one and two deaths not due to a cancer event respectively). The difference in the total number of disease free events between the two groups (i.e. 69 versus 55) is explicable by the n=14 difference in non-breast cancer malignancies between the two groups (i.e. 20 versus 6) which is likely a chance event.

This trial showed no difference in disease free survival, overall survival or recurrence. The 5-year disease-free survival, the primary endpoint, was 87.8% (95% CI 84.4–91.2) in the group without axillary dissection and 84.4% (80.7–88.1) in the group with axillary dissection (log-rank p=0.16) The HR for no axillary dissection vs axillary dissection was 0.78, 95% CI 0.55–1.11, non-inferiority p=0.0042, in favour of no axillary dissection). This improvement in the HR in favour of the group omitted ALND, which is surprising, is explicable by the difference in non-breast cancer malignancies (n=14)
Inclusion of ‘events’ which are not biologically related to the therapeutic intervention should by chance result in the same number of events in both arms – which will be a bias towards supporting non-inferiority. The authors provided no biological rationale for inclusion of some events such as contralateral breast cancer or second non-breast malignancies as events. Indeed from the perspective of loco-regional control the trial was significantly underpowered, with only a 16% power to detect non-inferiority in regional recurrence, 35% power to detect non-inferiority in either regional or distant recurrence and a 40% power to detect non-inferiority in ipsilateral breast, regional or distant recurrence. The study would have needed significantly longer follow up in order to be powered for these secondary endpoints.

In summary, while there is no evidence of benefit to ALND in patients with SNLB micrometastases, there are limitations to the data provided by IBCSG 23-01. A number of retrospective studies had previously reported that patients with occult lymph node metastases had similar outcome to patients with lymph node negative disease. Many experts therefore already believed that treatment of very low level lymph node involvement did not seem beneficial. In this respect IBCSG 23-01 was not practice-changing in the manner that the Z0011 study has been as the concept of leaving very low burden and low risk disease untreated by further regional therapy was not a significant leap.

4. Question 3: Is there a need for ALND in patients with SLNB macrometastasis? (ACOSOG Z0011 trial, ClinicalTrials.gov NCT00003855)

ACOSOG Z00112,6 was a multicentre trial for which targeted enrolment was 1900 women at least 18 years old undergoing breast conservation therapy who had clinical T1 or T2, N0, M0 breast cancer; one or two positive SLNs; and an Eastern Cooperative Oncology Group/Zubrod functional status =2, with final analysis planned after 500 deaths.

Patients with a positive SLNB were randomized to undergo either ALND or no further local treatment. The primary endpoint was overall survival with a non-inferiority margin of a hazard ratio of less than 1.3 indicating that SLND alone was non-inferior to ALND. 115 sites were opened enrolling patients from May 1999 to December 2004 – i.e. an expected average for each site of 16.5 patients. Over the 66 months the study was open this would have equated to an average of 3 patients per site per year.

As reported by the authors none of the planned interim analyses were performed before the study was closed based on the recommendation of the data and safety monitoring committee. Their recommendation was because “the study had accrued participants slower than projected, and both patient groups’ overall disease recurrence or death rate was much lower than expected”6. At this point 891 patients had been entered into the trial: 445 patients randomised to ALND and 446 randomised to SLND alone. This equated to 7.7 patients per site and for the 66 months the trial was open 1.4 patients per site per year. Since a study usually requires more patients to show non-inferiority than to show superiority, it was surprising that when the DMSC closed the study early based on a low death rate, that with less than half the initial targeted number of patients Z0011 should still be reported to show non-inferiority. Indeed in their early publication of the study, reporting on the morbidity between the
two arms of the study, the authors themselves stated that with early closure to recruitment by the DMSC “These developments may have impaired the ability of trial Z0011 to fulfil its primary objective, but we were still able to use the data to compare the two groups’ morbidities.”

On review of Z0011 there are a number of critical limitations including concerns regarding i) the relevance of some of the events contributing to its primary endpoint, ii) the final power of the study, iii) inclusion of ineligible patients, iv) baseline differences between groups and v) major protocol violations. Given the practice-changing nature of the Z0011 study all of these critical limitations will be discussed below.

4.1 Study design and endpoints
Firstly attention must be brought to the outcome criteria stated by the study for non-inferiority of SLNB alone. Overall survival was the primary end point, with a non-inferiority of less than 1.3 indicating that SLND alone is non-inferior to ALND. This means that by the omission of ALND for patients with 1 or 2 axillary lymph nodes that a survival difference of up to 30% more would be accepted as ‘confirming’ that SLNB was non-inferior. This is a very large margin to regard as an acceptable hazard ratio for a group of patients with only one or two positive nodes who were otherwise receiving optimal anti-cancer care including adjuvant systemic therapies. Indeed even systemic adjuvant therapy with tamoxifen does not provide a 30% difference in survival. Secondly given this very large hazard ratio it is even more surprising that when the study was closed early with less than half the patients recruited that the non-inferiority goal was still met. Even more surprising is the fact that the HR rather than being close to 1.3 was actually 0.79, implying that SLNB alone resulted in 21% improvement in survival over ALND. This counter-intuitive result will be discussed further below.

4.2 Recruitment rate and power
As noted above the study suffered from a very low recruitment rate (1.4 patients per site per year) and closed early due to a lower than expected death rate. This extremely low recruitment rate of patients with such a common condition raises questions as to i) the volume of breast cancer patients treated at these sites, ii) whether clinicians involved at participating centres were concerned about the possibility of patients not receiving treatment and thus refusing to enter them into the trial, and/or iii) only allowing a selected group of patients with better prognostic factors to be recruited.

The low recruitment rate, coupled with the much lower than expected rates of death, led to the investigators not reaching the statistical power they planned to. The a priori power calculation performed by the investigators required 1900 patients to be recruited, with 500 deaths in order to have 90% power to confirm non-inferiority. However having enrolled only 891 patients, of which only 94 died, led the study to have a much lower statistical power at around 35%.

The trial was also hampered by almost a fifth, (166/856, 19.4%), of patients being lost to follow-up. In view of the study’s low recruitment rate, it would have been even more imperative to have complete follow up for the patients that were recruited in order to maintain its required power. There is no evidence of a difference in the
distribution of patients dropping out between groups (22% in the ALND group vs 17% in the SLND group, p= 0.068) however this level of lost to follow-up at 5 years can only have further weakened the power of the study and its results. It is highly doubtful that a phase 3 registration trial of a new breast cancer drug would even be accepted for publication if 20% of its primary outcome data was missing due to patients being lost to follow-up, never mind being recommended for a change in standard of care.

4.3 Inclusion of ineligible patients
The Z0011 protocol deemed that women were ineligible if they had “3 or more positive SLNs, matted nodes, or gross extra-nodal disease, or if the patients had received neoadjuvant hormone or chemotherapy”. Despite this the authors also stated in the study publication “However, of the 891 registered patients, 287 were registered pre-SLND and assigned to treatment after intraoperative documentation of SLN metastases. Patients in this group subsequently found to have 3 or more tumour-involved lymph nodes were included in the analysis.” In other words although the study was only protocolled to include patients with 1 or 2 positive lymph nodes, the investigators knowingly included patients with 3 or more lymph nodes in the study and indeed included them in the analyses. This is despite having 3 or more positive lymph nodes being an ineligibility criteria. These ineligible patients accounted for 3.7% of patients in the SLN arm.

Furthermore in regards to ineligible patients the authors state “The 103 ineligible patients were included in the analyses reported here.” With 103/891 patients ineligible this represents 11.6% of patients recruited. Despite being ineligible, these patients were still included in the study analysis. At the same time the paper does not provide a description of the reasons why these 103 patients were deemed ineligible.

4.4 Baseline differences between groups
There is also concern over differences in the baseline lymph node status of patients between the two arms, with significantly more patients with a heavier burden of nodal disease being in the ALND group (See table 1). This was not commented on in the paper, instead the authors state “Disease characteristics at baseline were well balanced between the two groups.”

Firstly, the inclusion by the authors of a large number of patients with micrometastases and isolated tumour cells could only weaken the power of Z0011 in terms of its primary endpoint – i.e. overall survival – since this would reduce the number of deaths below that expected from patients with 1 or 2 macro-metastases positive lymph nodes. Indeed the low death rate was precisely why the DMSC closed the study early to recruitment.

However secondly, despite the authors stating that disease characteristics were well balanced between the two arms, this requires closer examination. The precise numbers given in the paper were 137 SLNB micrometastases/365 patients (37.5%) in the ALND arm versus 164/366 (44.8%) in the SLN alone arm. The authors simply state (p=.05) without any description of a statistical test or test value. Using a straightforward Chi square test (two tailed) the \( X^2 \) value = 3.99, with 1 degree of freedom; p=0.046. It appears that this p value was rounded up to 0.05 when quoted in
the paper, and in so doing implying this was of borderline significance between the two study arms.

This significant difference in micro-metastases in favour of the SLNB arm means that the SLNB group had a better prognosis at baseline and this must be considered as the most likely reason why the SLNB arm unexpectedly did better than the ALND arm. We say “unexpectedly” since the investigators of Z0011 project that SLNB would be worse than ALND. The fact the hazard ratio was 0.79 (95% CI 0.56-1.11) implies either that SLNB was ‘beneficial’ for the patients (alternatively that ALND was bad for them) or that the SLNB patients had a better prognosis from the outset. Given the significant difference in micro-metastases in the SLNB group the onus must be on the investigators to explain why SLNB is advantageous rather than Z0011 results more likely being due imbalance in baseline characteristics in favour of the SLNB group.

Another potentially confounding factor is that there was also a significant difference in the number of missing records of lymph node status between the two trial arms - as reported in table 1 of the Z0011 paper.\(^2\) (Table 1).

Comparing missing & known data in each of the two arms, 77/429 and 21/436, gives a $X^2 = 35.8$ 1 df; \(p < 0.001\) (i.e. highly significantly different between the two study arms). The authors do not comment on the potential effect this imbalance in missing data might have had in skewing the primary endpoint results.

<table>
<thead>
<tr>
<th>LN Mets</th>
<th>ALND (n=420)</th>
<th>SLNB alone (n=436)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4 (1.2)</td>
<td>29 (7.0)</td>
</tr>
<tr>
<td>1</td>
<td>109 (58.0)</td>
<td>295 (71.1)</td>
</tr>
<tr>
<td>2</td>
<td>66 (19.8)</td>
<td>76 (18.3)</td>
</tr>
<tr>
<td>3</td>
<td>25 (7.3)</td>
<td>11 (2.7)</td>
</tr>
<tr>
<td>≥4</td>
<td>47 (13.7)</td>
<td>4 (1.0)</td>
</tr>
<tr>
<td>Missing</td>
<td>77</td>
<td>21</td>
</tr>
</tbody>
</table>

### 4.5 Protocol violations

The study reports major protocol violations in relation to the axillary surgery performed. 32/420 (7.6%) patients randomized to ALND had SLN while 11/436 (2.5%) randomized to SLN had ALND and were subsequently included in all analyses. It is difficult to imagine that this was due to inadvertent error since this degree of performing the ‘wrong’ operation would not be acceptable in any competent healthcare facility. It has to therefore be assumed, until explained otherwise, that either the patients withdrew consent (in which case the patient should have been withdrawn from the study before surgery) or the surgeon performing the operation deliberately chose to perform an operation other than what the patient was randomized to receive - i.e. a protocol violation.
In terms of protocol violations particular concern also relates to the radiotherapy (RT) given in the context of the trial. This was reported separately by some of the original authors in a following paper and commented on by Zellars. On review of the RT data only 605/856 (70.7%) patients had a study case record form available, meaning that 29.3% of patients had missing RT case record forms. This has to be seriously concerning in terms of the completeness of data and the data analyses.

Jagsi et al state that they “attempted to obtain detailed radiation records for 791 patients [i.e. their individual hospital records]. Detailed RT records were received for central review for 228 of these patients (29%). Of these, 104 were from the ALND arm (26.7% of 389 patients randomly assigned to this arm), and 124 were from the SLND arm (30.7% of 404 patients randomly assigned to this arm). Among the 228 patients for whom detailed RT records were available, 138 had documentation of three-dimensional treatment planning.” It would appear that only for 138/228 (60.5%) patient records reviewed by Jagsi there were three dimensional RT planning which means that for 39.5% there was no documentation of RT planning. Jagsi also noted “Patients with RT information available did not differ significantly from those for whom it was not.” suggesting that this degree of missing data on each patients individual case records for three dimensional planning was across the whole study population – this alongside the fact that 29.3% of Z0011 study RT case record forms were missing.

Jagsi reported on the clinical characteristics of the 228 patients for whom detailed RT records were available. Within this subgroup, patients in the SLND arm had fewer lymph node metastases (P<0.001).

An additional concern is the number of radiotherapy treatment protocol violations. Large number of patients received high tangential fields, nodal radiotherapy or a posterior axillary field boost. As regards high tangential field the percentage of patients in the subgroup reviewed by Jagsi et al was “50% of patients (33 of 66) randomly assigned to the ALND arm and 52.6% of patients (40 of 76) randomly assigned to the SLNB arm.” The argument has been proposed by the proponents of Z0011 that the use of high tangential fields was balanced between the arms and therefore would not affect the primary outcome. However the value of RT after ALND is not the same as RT given to an untreated axilla. This is therefore not only a major protocol violation but one which can significantly affect the study endpoints.

Of the 228 patients whose records Jagsi et al reviewed, 43 (18.9%) received directed regional nodal RT using ≥ three fields: 22 in the ALND arm and 21 in the SLNB arm. Giving RT to the axilla post ALND and post SLNB are quite different in terms of the therapeutic effect the RT might have on disease recurrence. Irrespective of the difference in effect between the two arms they represent a further 18.9% of patients with protocol violations. Furthermore Jagsi reported that “Those receiving nodal RT had a greater number of lymph nodes involved (P<0.001) than those who did not [receive radiotherapy].” This shows that the use of nodal RT was not a random event but was associated with the degree of involvement of the axillary lymph nodes and therefore almost certainly due to physician selection. Again this bias to treat more disease burdened axilla by nodal RT is likely to have had greater effect in the SLNB arm (i.e. effectively an otherwise untreated axilla) compared to the ALND arm. Furthermore not only was the axilla treated in 18.9% of patients and the posterior
axillary field boosted in 57% but 16.5% of patients in the SLNB group reviewed by Jagsi also received supraclavicular fossa (SCF) RT. The use of SCF RT in 16.5% of patients where 44.8% of SLNB patients had micro-metastases and the remainder had 1-2 axillary lymph node macro-metastases is difficult to explain on the clinical criteria described in the Z0011 publication.

All of the above RT protocol violations meant that large numbers of patients in the SNLB group actually received effective local treatment to the axilla in the form of radiotherapy, thereby reducing any observable benefit from ALND versus SLNB alone.

4.6 Summary
In summary, Z0011 has several critical limitations/flaws in the study and as a result the report that SNLB alone is not inferior to ALND in patients with 1-2 SLNB macro-metastases is critically flawed. Furthermore the results of Z0011 do not easily match with the data from MA20 trial\textsuperscript{11,12} where regional RT in addition to ALND improved disease outcome over ALND alone (see comments on MA 20 below).

5. Question 4: Is axillary radiotherapy equivalent to ALND in patients with SNLB macrometastasis? (AMAROS trial, ClinicalTrials.gov NCT 00014612)

The After Mapping of the Axilla: Radiotherapy or Surgery? (AMAROS) phase 3 trial\textsuperscript{13} randomized 1425 patients with a positive SLNB to either undergo ALND or axillary radiotherapy. The main objective of the trial was to show non-inferiority of the radiotherapy arm (RT) compared with the ALND treatment arm with respect to axillary recurrence-free rate in sentinel node–positive patients.

The primary endpoint of this study therefore was axillary recurrence free survival. As noted above this is different from Z0011 trial which included events such as contralateral breast cancers and second non-breast cancer malignancies.

AMAROS appears to have been carried out to a high standard but like the IBCSG 23-01 study it was underpowered in terms of the number of patients to be recruited. AMAROS was also limited to patients with tumours of low risk of recurrence – as reflected in the entry criteria and baseline tumour characteristics. For example tumour size on entry was to be between 0.5-3.0 cm and had to be clinically node negative however the median tumour size was 18mm (IQR 13-23) for the patients randomised to radiotherapy and 62% had a pre-operative ultrasound of the axilla. Following SLNB, 93%-94% of patients had only 1 or 2 lymph nodes positive with 34%-37% being involved with only micro-metastases or isolated tumour cells. Again this high proportion of micro-metastases and isolated tumour cells would be contributory to the low recurrence rate and thus the low power in the AMAROS study.

Despite this low risk population with its low recurrence rate, it should be noted that the recurrence rate in the RT arm was double the ALND (1.03% and 0.54% respectively) with median follow up of 6.1 years. While the absolute difference is small between the two low risk arms in the AMAROS study, clinicians should be wary about extending the idea of equivalence of RT and ALND to patients at higher risk of recurrence. In addition, it may be incorrect to conclude that axillary radiotherapy gives less side effects than ALND. Firstly, although level 1 & 2 ALND was accepted, level 3 was strongly encouraged, which may not be the case today
within an intermediate risk population. Secondly, it is premature to conclude that lymphoedema rates are less with radiotherapy at 5 years, as late normal tissue side effects are well known to continue to develop more than 5 after completion of treatment. Thirdly, post-hoc analysis suggested that axillary radiotherapy patients had more problems with shoulder movement than ALND, which demonstrates that lymphoedema is not the only endpoint to take into account. Finally, with modern CT planning, it is possible that the radiotherapy coverage to axillary nodal groups will be “better” and therefore the patient could receive an effectively higher dose than within AMAROS, with the potential for more toxicity. Therefore, it is essential that modern vessel-based contouring guidelines such as the ESTRO consensus14 are used to accurately define the nodal regions and guide radiotherapy.

In summary, while the AMAROS study showed no evidence of a difference in recurrence, it was underpowered due to the low number of recurrences which was also a reflection of the low risk population including a third of patients having only micrometastases or isolated tumour cells in the SLNs.

6. **Question 5: Does regional nodal irradiation (RNI) add to whole breast irradiation (WBI) after BCS for women with node positive or high risk node negative breast cancer treated with adjuvant systemic therapy? (MA20 trial, ClinicalTrials.gov NCT00005957)**

Patients could enter MA20 if their tumour was node positive (1-3 positive LNs) or was high risk node negative – the latter being defined as either tumour > 5cm diameter or tumour > 2cm diameter and less than 10 axillary nodes removed plus either being oestrogen receptor negative, grade 3 or display evidence of lymphovascular invasion. Only 10% of the 1832 patients randomised were node negative. 85% of the patients had 1-3 positive lymph nodes. Therefore it appears that the majority of patients in MA 20 cover the same 1-2 lymph node positive disease to which Z0011 is deemed to be relevant.

The primary endpoint of MA 20 was overall survival and at a planned interim analysis after a median of 62 months the 5 year survival of the WBI only was 90.7% versus 92.3% for WBI + RNI (HR=0.76; p=0.07). For the secondary endpoints of DFS, the results were 84% for WBI only and 89.7% for WBI + RNI (HR=0.67; p=0.003) and for distant DFS 87% for WBI only and 92.4% for WBI + RNI (HR=0.64; p=0.002)11.

While the study did not meet its primary endpoint of overall survival (p=0.07) it is clear that treatment of the axilla with RNI adds significantly to WBI and axillary surgery (ALND in >85% of patients) in terms of disease control, both DFS and distant DFS. In addition while loco-regional recurrence may not always lead to death, distant recurrence is currently an incurable disease.

A subsequent publication reported the final analysis in which the median follow-up was 9.5 years and the 10 year survival rates were 81.8% in the WBI only (control) group and 82.8% in the WBI+RNI group (HR for death in the nodal-irradiation group as compared with the control group, 0.91; 95% confidence interval [CI], 0.72 to 1.13; P = 0.38). There was a prespecified subgroup analysis looking at patients with ER-
negative tumours. The 10-year overall survival rate was higher in the WBI+RNI group compared with the WBI only group (81.3% vs. 73.9%), a difference that approached statistical significance (hazard ratio, 0.69; 95% CI, 0.47 to 1.00; P = 0.05). There was no significant difference detected in breast-cancer mortality, with 10-year rates of 10.3% in the WBI+RNI group versus 12.3% in the WBI only group (HR=0.80; 95% CI, 0.61 to 1.05; P = 0.11) – this represented a difference of 20% in breast cancer mortality, which was not statistically significant. The causes of non-breast-cancer deaths were similar in the two groups.

Comparing WBI+RNI group with the WBI only group the 10 year rates for the secondary endpoints were as follows:- the rate of disease-free survival was higher in the WBI+RNI group than in the RNI only group, with 10-year rates of 82.0% and 77.0%, respectively (HR=0.76; 95% CI, 0.61 to 0.94; P = 0.01); the isolated locoregional disease-free survival were 95.2% in the WBI+RNI group versus 92.2% in the WBI only group (HR=0.59; 95% CI, 0.39 to 0.88; P = 0.009); the rates of distant disease-free survival was higher in the WBI+RNI group than in the WBI only group, with 10 years of 86.3% and 82.4% respectively (HR=0.76; 95% CI, 0.60 to 0.97; P = 0.03)\(^{(10)}\).

The results from MA 20 are important in relation to interpreting Z0011, since the Z0011 investigators have claimed that patients with 1-2 positive axillary lymph nodes do not need any further regional treatment whereas MA20 shows that RNI even in addition to ALND can result in improved disease control. There is now further support for this argument with the publication of the similar EORTC\(^{(15)}\) and Danish Breast Cancer Co-operative trials,\(^{(16)}\) which showed an improvement in DFS and also OS respectively.

### 7. Overall Summary

The findings of NSABP B32 were that in the presence of a negative sentinel lymph node that no further regional treatment was non-inferior to an ALND – although it should be noted that the analysis yielded an unadjusted hazard ratio (HR) of 1·20 (95% CI 0·96–1·50) – implying a 20% difference in overall survival in favour of ALND. In a lymph node negative population, particularly with biologically good prognostic tumours, the absolute risk of death will be low and in this situation a 20% relative difference would be a very low absolute benefit.

The findings of IBCSG 23-01 suggest that in the presence of micro-metastases that no further regional treatment was non-inferior to an ALND – although it should be noted that this study was severely underpowered to show non-inferiority in regional recurrence, distant recurrence or overall survival.

The findings of Z0011 have to be critically re-assessed in view of the significant flaws in the data presented (e.g. imbalance in baseline characteristics, imbalance in missing follow-up data, inclusion of ineligible patients in the results analyses, major protocol violations) as well as the subsequent reports on missing RT data records. All of the preceding problems identify critical limitations in using the reported results from Z0011 for changing the standard of care for patients with 1 or 2 lymph node positive disease who would most commonly fall into an intermediate risk category but could even fall into a high risk category (e.g grade 3, T2 tumour). Furthermore the results of MA 20, which does not have the critical limitations of Z0011 presents a contrary message that RNI for 1-3 nodes provides a disease outcome benefit.
This finding is consistent with previous reports from the EBCTCG Overview of the value of regional RT in patients with less than 4 positive axillary lymph nodes.\textsuperscript{17} One criticism of the EBCTCG overview has been that these data come from an era in which systemic therapies were not as effective as currently used. The importance of this may be debatable but this criticism cannot be made of the recent MA 20 trial where current systemic therapy regimens were used.

In addition to the above, the results of the SUPREMO\textsuperscript{18} Phase 3 RCT will be of importance in addressing and answering the need (or not) for additional regional radiotherapy in the treatment of low burden axillary disease. In SUPREMO 1600 patients who had mastectomy with either less than 4 LNs positive or high risk node negative disease after undergoing ALND were then randomized to nodal radiotherapy or not. The results of this study are awaited.

The findings of AMAROS suggest that RT may be non-inferior to ALND in patients with low risk primary tumours – again it should be noted that this study was also underpowered.

Finally the MA 20 study, which failed to achieve its primary endpoint on overall survival, did report consistently significant improvements in disease outcome in terms of DFS and distant DFS; the latter can reasonably be expected to be a forerunner of overall survival.

We would note that there are good examples in breast cancer where less aggressive treatment has been shown not to adversely affect outcome.\textsuperscript{19} Z0011 was a groundbreaking idea but given the limitations of Z0011 the hypothesis that we can omit regional treatment (ie ALND &/or radiotherapy) for patients with 1-2 macrometastases still requires testing in a prospective randomized trial. It is also worth considering the implication of extra-nodal invasion in this context and further studies are required.

Other avenues of future study include the concept of targeted lymph node clearance in the context of neoadjuvant chemotherapy which show early promise.\textsuperscript{20}

POSNOC (ClinicalTrials.gov NCT02401685) is an ongoing Phase 3 RCT which will randomise 1900 patients with 1 or 2 macro-metastases to either standard axillary treatment (ALND or RT) versus no further regional treatment.\textsuperscript{21} The Z0011 study only recruited patients who had breast conserving surgery with post-operative intact breast irradiation. As noted above it appears a number of the patients in Z0011 also received radiotherapy to the low axilla through the selection by the radiation oncologist of high tangential fields. POSNOC will include both patients undergoing breast conserving surgery or mastectomy who are both clinically and radiologically node-negative preoperatively, with the primary endpoint of axillary recurrence at 5yrs. The majority of patients will not require radiotherapy post-mastectomy; this will remove a potential confounding problem compared to those patients receiving breast conserving surgery (ie high tangential fields) POSNOC will hopefully provide an answer to this important question of whether patients with known low burden disease in the axilla can omit regional treatment(s) and be treated only with systemic adjuvant therapiess. The stratification by breast surgery (ie breast conserving surgery or mastectomy) will help tease out whether radiotherapy to the breast with inadvertent
axillary radiotherapy due to high tangential fields might have been partly responsible for the results of Z0011.

When planning further trials, it is worth considering the differences in the method used to determine cN0 prior to study. NSABP-32 simply states “clinically node negative”, IBCSG23-01, Z00011 and the AMAROS excluded those with palpable lymph nodes. POSNOC requires a pre-operative axillary ultrasound.

Similarly, there are important differences in the histological techniques in the trials discussed above, with NSABP-32 and IBCSG23-01 using routine haematoxylin and eosin (H&E), allowing Immunohistochemistry (IHC) only to confirm suspicious findings on H&E. Z0011 did not allow IHC, but allowed H&E, frozen sections or touch preparation. AMAROS subjected all H&E negative tissue to IHC. MA20 did not comment on its histopathology in its methods and POSNOC will allow IHC for selective characterisation of suspicious tissue.

Until the POSNOC study reports, it is the view of these authors that Z0011 does not provide a level and quality of evidence which would support no axillary treatment of 1-2 macro-metastases as the standard of care, particularly in the light of MA20 where more regional therapy appears to give better disease control.

**Conflict of interest**

The authors have no conflicts of interest to declare.

**References**


**Source of funding**

No funding was sought to support this paper.