Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial


Summary
Background Sentinel-lymph-node (SLN) surgery was designed to minimise the side-effects of lymph-node surgery but still offer outcomes equivalent to axillary-lymph-node dissection (ALND). The aims of National Surgical Adjuvant Breast and Bowel Project Breast and Bowel Project (NSABP) trial B-32 were to establish whether SLN resection in patients with breast cancer achieves the same survival and regional control as ALND, but with fewer side-effects.

Methods NSABP B-32 was a randomised controlled phase 3 trial done at 80 centres in Canada and the USA between May 1, 1999, and Feb 29, 2004. Women with invasive breast cancer were randomly assigned to either SLN resection plus ALND (group 1) or to SLN resection alone with ALND only if the SLNs were positive (group 2). Randomisation assignment was done at the NSABP Biostatistical Center (Pittsburgh, PA, USA) with the biased coin minimisation approach in an allocation ratio of 1:1. Stratification variables were age at entry (<49 years, ≥50 years), clinical tumour size (≤2·0 cm, 2·1–4·0 cm, ≥4·1 cm), and surgical plan (lumpectomy, mastectomy). SLN resection was done with a blue dye and radioactive tracer. Outcome analyses were done in patients who were assessed as having pathologically negative sentinel nodes and for whom follow-up data were available. The primary endpoint was overall survival. Analyses were done on an intention-to-treat basis. All deaths, irrespective of cause, were included. The mean time on study for the SLN-negative patients with follow-up information was 95·6 months (range 70·1–126·7). This study is registered with ClinicalTrials.gov, number NCT00003830.

Findings 5611 women were randomly assigned to the treatment groups, 3989 had pathologically negative SLN. 309 deaths were reported in the 3986 SLN-negative patients with follow-up information: 140 of 1975 patients in group 1 and 169 of 1976 patients in group 2. Log-rank comparison of overall survival in groups 1 and 2 yielded an unadjusted hazard ratio (HR) of 1·20 (95% CI 0·96–1·50; p=0·12). 8-year Kaplan-Meier estimates for overall survival were 91·8% (95% CI 90·4–93·3) in group 1 and 90·3% (88·8–91·8) in group 2. Treatment comparisons for disease-free survival yielded an unadjusted HR of 1·05 (95% CI 0·90–1·22; p=0·54). 8-year Kaplan-Meier estimates for disease-free survival were 82·4% (80·5–84·4) in group 1 and 81·5% (79·6–83·4) in group 2. There were eight regional-node recurrences as first events in group 1 and 14 in group 2 (p=0·22). Patients are continuing follow-up for longer-term assessment of survival and regional control. The most common adverse events were allergic reactions, mostly related to the administration of the blue dye.

Interpretation Overall survival, disease-free survival, and regional control were statistically equivalent between groups. When the SLN is negative, SLN surgery alone with no further ALND is an appropriate, safe, and effective therapy for breast cancer patients with clinically negative lymph nodes.

Funding US Public Health Service, National Cancer Institute, and Department of Health and Human Services.

Introduction Axillary-lymph-node dissection (ALND) for breast cancer is a procedure designed to maximise survival and regional control and to establish nodal classification. However, this procedure is associated with short-term and long-term side-effects in a substantial number of patients. Sentinel-lymph-node (SLN) resection was designed to minimise the side-effects of lymph-node surgery, but still offer outcomes equivalent to ALND.

The National Surgical Adjuvant Breast and Bowel Project (NSABP) B-32 trial was therefore designed to establish whether SLN resection achieves the same therapeutic outcomes as ALND, but with fewer side-effects. Primary endpoints of the B-32 trial were survival, regional control, and morbidity. Technical outcomes and assessment of the training methods for this trial have been reported previously. Patient-reported outcomes and morbidity related to range of motion, oedema, pain, and sensory defects have also been reported. An ancillary pathology study assessing survival according to the detection of occult node metastases in 3887 SLN-negative patients has also been completed and will be reported separately.

We report the primary outcome survival data from the NSABP B-32 trial, which aimed to establish whether SLN resection achieves the same therapeutic outcomes as ALND, but with fewer side-effects.
resection in patients with SLN-negative breast cancer achieves the same survival and regional control as ALND, but with fewer side-effects. The data presented are based on a randomised controlled trial that used standardised surgical and pathological methods to ensure that the primary outcomes were as comparable as possible between treatment groups.

Methods

Patients

Between May 1, 1999, and Feb 29, 2004, a randomised controlled phase 3 trial was done at 80 institutions in Canada and the USA. Women with invasive breast cancer and clinically negative nodes were randomly assigned to SLN resection plus ALND (group 1) or to SLN resection alone with ALND only if SLNs were positive (group 2). Patients in both groups with pathologically negative SLNs were monitored at 4–6 month intervals for overall survival, disease-free survival, and regional control. Monitoring procedures included bilateral mammogram every 12 months and history and physical examination every 6 months up to year 3 and every 12 months thereafter. Patients in group 1 with negative SLNs and positive non-SLNs were included in group 1 as SLN-negative. All outcome data reported here are for SLN-negative patients.

NSABP B-32 was done after approval from local institutional review boards and in accordance with assurances filed with and approved by the US Department of Health and Human Services. Informed consent was obtained from each participant in this study.

Randomisation and masking

Patients were randomly assigned in a 1:1 ratio at the NSABP Biostatistical Center (Pittsburgh, PA, USA). Stratified randomisation was done by use of a biased coin minimisation approach. The algorithm was used to randomly generate a treatment assignment at the time each individual was entered into the study. Stratification factors included age at entry (≤49 years, ≥50 years), clinical tumour size (≤2.0 cm, 2.1–4.0 cm, ≥4.1 cm), and surgical treatment plan (lumpectomy, mastectomy). Masking of type of nodal dissection was not possible because of the nature of the procedures.

Procedures

Details of training and quality control have been previously reported. Surgeons and pathologists were required to follow specific protocols for SLN surgery, labelling of lymph nodes, and for pathological analysis of the lymph nodes. Performance audits documented excellent adherence to protocol.

99mTechnetium-sulfur colloid (99mTc-TSC) was injected into the breast around the tumour and intradermally over the tumour from 30 min to 8 h before surgery. Isosulfan blue was injected into the breast around the tumour 5 min before incision. Lymph nodes that were radioactive, blue, or clinically positive were judged to be sentinel nodes. If a non-SLN was removed during an SLN procedure it was submitted to pathology separate from the sentinel nodes and labelled as a non-SLN. SLNs from both groups 1 and 2 were assessed postoperatively with routine stains at about 2 mm intervals through the node. Immunohistochemistry was not permitted, except for confirmation of suspicious findings on routine haematoxylin and eosin stains. Additionally, SLNs from group 2 were assessed intraoperatively with cytology.

Statistical analysis

The primary endpoint, overall survival, included all deaths from the time of randomisation. Disease-free survival events included local, regional, or distant breast cancer recurrences; second cancers (in the opposite breast and non-breast cancer); and all deaths from the time of randomisation. Analyses were done according to the random assignment of patients in those with pathologically negative sentinel nodes and follow-up information. For patients who withdrew consent for further follow-up after they were randomly assigned, any events up to the time of withdrawal of consent were included. The B-32 trial was powered to test a difference in survival of 2% between the two groups for the SLN-negative patients at 5 years. Formal interim endpoint analyses were presented to an external data monitoring committee after 71, 148, and 242 deaths had been reported. In all these interim analyses, the committee recommended continuation of the trial without divulging early results. 300 deaths were needed to trigger the overall survival analysis; this number of events was met in the trial.
December, 2009. Therefore, the analyses here are based on data up to Dec 31, 2009.

Simple log-rank tests and Cox proportional hazard models were used to make formal inferences about group comparisons, and Kaplan-Meier curves were used to quantify the values of overall survival and disease-free survival over time.9 In the Cox regression analyses, adjustments were made for the stratification variables.10 Tests of the validity of the proportionality assumption were done via the method proposed by Grambsch and Therneau.11 Cox models were also used to establish if significant treatment by stratification variable interactions existed with respect to the endpoints.12 Site-specific failure rates were calculated with cumulative incidence curves.13,14 These estimates appropriately adjust for competing risks of failure.13 p values for treatment comparisons of cumulative incidence curves were obtained with cause-specific hazard rates.15 In the forest plots used to display subset analyses, because the comparisons involve stratification variables, the hazard ratios (HRs) show adjusted treatment comparisons. 95% CIs are reported and all CIs and p values are two-sided. The α-level for statistical significance is 0·05. Statistical analyses were done with SAS version 8.4 and R version 2.8.0. This trial is registered with ClinicalTrials.gov, number NCT00003830.

Role of the funding source
The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. SJA, TA, and JPC had full access to the raw data in the study and had final responsibility for the decision to submit for publication.

Results
Figure 1 shows the trial profile. Women with clinically negative axillary lymph nodes, as assessed by physical examination, were randomly assigned to group 1 or group 2. 3989 patients (71·1%) of 5611 were SLN-negative, of whom 3986 (99·9%) had follow-up data. Random assignment of patients to the two treatment groups was balanced according to age, clinical tumour size, and surgical treatment plan (table 1).

The use of systemic adjuvant therapy was well balanced, with 85·1% (1680 of 1975 patients) of group 1 receiving therapy and 84·2% (1694 of 2011) of group 2. Radiation therapy was also well balanced, with 81·9% (1618 of 1975) of group 1 receiving therapy and 82·0% (1650 of 2011) of group 2.

The mean time on study for the SLN-negative patients with follow-up information was 95·6 months (range 70·1–126·7). Among the 3986 women with follow-up data, the deaths of 309 patients were reported (140 of 1975 in group 1 and 169 of 2011 in group 2). The mean yearly mortality across the two groups was 1·12% (95% CI 1·00–1·24; 1·02% in group 1 [0·85–1·19] and 1·22% in group 2 [1·03–1·40]). Of the 309 deaths, 109 happened after the first event was a breast cancer recurrence (50 in group 1 and 59 in group 2). Of these 109 deaths, 15 were after a local recurrence (eight in group 1 and seven in group 2), ten after regional recurrence (three in group 1 and seven in group 2), and 84 after systemic recurrence (39 in group 1 and 45 in group 2). There were seven deaths after a new
contralateral breast cancer (five in group 1 and two in group 2). There were 84 deaths after the first event was a second non-breast cancer (32 in group 1 and 52 in group 2) without a recurrence of their breast cancer or a second cancer. In the subset of patients in group 1 who had positive axillary nodes and negative sentinel nodes, five of 75 died (0·99% mean yearly mortality; 95% CI 0·33–2·40).

Log-rank comparison of overall survival in group 1 and group 2 yielded an unadjusted HR of 1·20 (95% CI 0·96–1·50; p=0·12). Overall, mortality in group 1 was lower than that noted in group 2. However, the two groups were statistically equivalent since the 95% CI for the mortality HR crossed 1. Cox proportional hazard analyses adjusting for stratification variables yielded results very similar to the unadjusted results (HR 1·19, 95% CI 0·95–1·49; p=0·13). The test for the interaction of treatment with all stratification variables combined for overall survival yielded a non-significant result (p=0·25). Furthermore, none of the individual stratification variables had significant interactions with treatment.

5-year Kaplan-Meier estimates for overall survival were 96·4% (95% CI 95·6–97·2) in group 1 and 95·0% (94·0–96·0) in group 2; the 8-year estimates were 91·8% (90·4–93·3) for group 1 and 90·3% (88·8–91·8) for group 2 (figure 2).

Treatment comparisons for disease-free survival yielded an unadjusted HR of 1·05 (95% CI 0·90–1·22; p=0·54); the adjusted HR was 1·07 (0·90–1·22; p=0·57). Table 2 shows the location of first treatment failure. No substantial differences are evident across sites. The mean yearly event rate pooled across the two groups is 2·49% (95% CI 2·30–2·68; 2·43% [2·16–2·70] in group 1 and 2·55% [2·28–2·82] in group 2). In the subset of patients in group 1 who had positive axillary nodes but negative sentinel nodes, ten of 75 had events (2·06% mean annual event rate; 95% CI 0·78–3·33).

Figure 4 summarises the HRs and 95% CIs of the two groups for all sites of first treatment failures. No significant differences were recorded.

There were 54 local recurrences in group 1 and 49 in group 2 (p=0·55). 99 of the 103 local recurrences were ipsilateral breast tumour recurrences (51 in group 1 and 48 in group 2), three were in the chest wall (two in group 1 and one in group 2), and one (in group 1) was in the area of the surgical scar.

There were eight regional node recurrences as first events in group 1 and 14 in group 2 (p=0·22). Among the 22 regional events, ten were in the axilla (two in group 1 and eight in group 2), seven were in the supraclavicular area (three in group 1 and four in group 2), one (in group 1) in the parasternal region, one (in group 1) in the subclavicular area, and three in both local and regional areas (one in group 1 and two in group 2).

46 patients (0·8%) had adverse events in the form of allergic reactions: 24 were grade 1 reactions, nine grade 2, three grade 3, and ten grade 4. Most of these reactions were related to the blue dye. Also, 14 (0·5%) of 2788 patients of group 1 and 12 (0·4%) of 2800 of group 2 had grade 3 or greater surgery-related adverse events.

**Discussion**

Our trial shows that overall survival, disease-free survival, and regional control were all statistically equivalent in SLN-negative patients who had an ALND (group 1) or SLN surgery alone (group 2). The survival difference between the groups was less than 2% and any variation under that threshold is not significant. In a trial of this size we did not expect exact numerical duplication of events. In group 1, 75 patients had at least
one positive non-sentinel node and 95% of this subset (71 patients) were treated with systemic adjuvant therapy. The outcome of these 75 patients was not worse than that of the group as a whole (mean yearly mortality 0.99% [95% CI 0.32–2.30] vs 1.02% [0.85–1.19], respectively), even though they were node-positive. In group 2, a similar subset of non-SLN-positive patients might be expected; however, since their axillary nodal status was unknown, the proportion of potentially non-SLN-positive receiving adjuvant therapy was likely to be similar to the remaining group 2 patients (84–1%). This might have contributed slightly to the hazard ratio for overall survival being in favour, albeit non-significantly, of the group 1 patients. Also, after a second non-breast cancer there were 32 deaths in group 1 and 52 in group 2. Non-breast cancer deaths, randomly in favour of group 1, might have also contributed to survival differences between groups.

Disease-free survival was not different between the two treatment groups. Comparisons based on sites of first treatment failures also showed no significant differences across all sites. These data further confirm the similarity in outcomes between the two treatment groups.

Each treatment group had less than 1% regional recurrences as first events. Similar to several non-randomised reports, the B-32 results confirm the low rate of regional-node recurrences after SLN surgery. The trial also shows that in patients with negative SLNs the number of regional node recurrences does not differ significantly between patients who have axillary dissection or SLN resection only.

Previous B-32 results showed that patient-reported outcomes and morbidity related to range of motion, oedema, pain, and sensory defects are lower in the SLN group than in the ALND group also confirm other literature findings. SLN surgery is itself not without complications, however, and there is a small increase over baseline of extremity oedema and functional and neurological deficits.

Randomised trials have been instrumental in effecting changes in the surgical management of breast cancer. One of the last major surgical trials that led to the safe reduction of surgery, NSABP B-06, was first reported in 1985. The longstanding importance of this trial is shown by the current use of breast-conserving therapy as a major indicator of quality care. SLN surgery represents the next major step in reducing the extent of surgical procedures to treat breast cancer.

The design of these randomised trials for breast-conserving therapy and lymph-node-sparing surgery are similar, with goals to preserve tissue and decrease morbidity, but still achieve the same cancer control so survival is not adversely affected. The B-32 trial was designed so that a relatively small 2% difference in survival would be detected. This narrow difference in survival was chosen to ensure that reduced morbidity was not at the expense of reduced survival. This required a substantial number of patients to be accrued and is why the B-32 trial is the largest randomised surgical trial in breast cancer thus far (panel).

One measure of quality in trial design is the clarity of the goals. The primary outcomes of B-32 were clearly stated and the trial was monitored regularly by an independent data monitoring committee. Potential imprecision was possible because of the complexity of surgical and pathological procedures but this variation was controlled by a careful preregistration training programme that focused on protocol compliance. Additionally, extensive auditing of enrolled cases assessed 94 specific items per patient. Of the 224 surgeons audited, protocol compliance was excellent. The quality of the trial is further supported by the extent of follow-up information (99–90% of cases).

Other trials include the American College of Surgeons Oncology Group trial Z0011 that randomly assigned 891 patients with pathologically positive SLNs to ALND or SLN resection only. This study closed before meeting its accrual goals. A trial from Milan has reported data with 10 years of follow-up. The primary objective for the Milan trial was “the predictive power of the status of the sentinel node”. Survival data comparing patients who were treated only with SLN resection are not available. The primary outcomes from the ALMANAC and SNAC trials are morbidity, and these well designed trials do not address survival. The AMAROS trial is another well designed randomised trial comparing ALND to SLN resection and radiation therapy with no further ALND. The primary objectives for the AMAROS trial were local and regional control, and morbidity.

Data combined from the available randomised trials of ALND versus no ALND suggest a slight survival advantage to ALND. Survival has also been significantly associated with the number of nodes removed. SLN surgery is neither observation-only nor removal of suspicious nodes from a fixed anatomical location, but is the highly targeted removal of the lymph nodes that receive direct drainage from a solid tumour in the breast. The results from B-32 show that in the SLN-negative
population, any survival advantage of full ALND is fully mitigated by simply removing the SLNs.

Surgeons should continue to strive to optimise the methods of SLN surgery. For example, removal of only a single SLN increases the risk for false-negative SLN resection. Improving methods to validate that the nodes removed are in fact on the immediate drainage pathway from the cancer is important. Also, life-threatening anaphylactic reactions related to dyes happen in about 0·25–0·5% of cases. Genotoxicity of blue dyes in the form of DNA-strand breaks and increases in oxidative DNA lesions have been reported after very brief exposure to cells in vitro. Therefore, there are clearly unfinished areas of research in the specialty of SLN surgery.

In summary, NSABP B-32 results suggest that when the SLN is negative, SLN surgery alone with no further ALND is an appropriate, safe, and effective therapy for patients with breast cancer.

Contributors
DNK was the principal investigator. SJA was involved in the data preparation, data analysis, writing the report and its review. TJG was involved in protocol management and writing the report, and was the institutional principal investigator. AMB was involved in the development of the data forms, review of the report, and critique. SPH was the co-investigator in charge of surgical training for the trial. JFC was involved in data collection, interpretation of results, and review of the report, its editing, and its final approval. TA was involved in reviewing the report. DWL was involved in data collection and analysis, study design, and writing and editing the report. EFM was involved in study design and conduct, patient accrual, and preparation and approval of the report. LMJ was involved in patient accrual and review of the data and the report. TGF contributed numerous patients to the study and helped to review the report. RDN was involved in patient accrual to the B-32 trial. AR was involved in patient accrual, and review of the report and its final approval.

HMCS was involved in patient accrual and review of the report. NW provided administrative support and was involved in the concept and design of the study, data interpretation, and final approval.

Conflicts of interest
The authors declared no conflicts of interest.

Acknowledgments
This work was funded by PHS grants U10CA-12027, U10CA-37377, U10CA-69353, and U10CA-69974 (NSABP), P30 CA22415 (Vermont Cancer Center), SRO1CA74137 (DNK), and U10 CA44606 (AR). NCI Department of Health and Human Services. We thank Barbara C Good and Wendy I. Rea for editorial assistance, and Elaine Harper for data management.

References


