Surgeon Training, Protocol Compliance, and Technical Outcomes From Breast Cancer Sentinel Lymph Node Randomized Trial

David N. Krag, Takamaru Ashikaga, Seth P. Harlow, Joan M. Skelly, Thomas B. Julian, Ann M. Brown, Donald L. Weaver, Norman Wolmark; for the National Surgical Adjuvant Breast and Bowel Project

Background The National Surgical Adjuvant Breast and Bowel Project B-32 trial was designed to determine whether sentinel lymph node resection can achieve the same therapeutic outcomes as axillary lymph node resection but with fewer side effects and is one of the most carefully controlled and monitored randomized trials in the field of surgical oncology. We evaluated the relationship of surgeon trial preparation, protocol compliance audit, and technical outcomes.

Methods Preparation for this trial included a protocol manual, a site visit with key participants, an intraoperative session with the surgeon, and prerandomization documentation of protocol compliance. Training categories included surgeons who submitted material on five prerandomization surgeries and were trained by a core trainer (category 1) or by a site trainer (category 2). An expedited group (category 3) included surgeons with extensive experience who submitted material on one prerandomization surgery. At completion of training, surgeons could accrue patients. Two hundred twenty-four surgeons enrolled 4994 patients with breast cancer and were audited for 94 specific items in the following four categories: procedural, operative note, pathology report, and data entry. The relationship of training method; protocol compliance performance audit; and the technical outcomes of the sentinel lymph node resection rate, false-negative rate, and number of sentinel lymph nodes removed was determined. All statistical tests were two-sided.

Results The overall sentinel lymph node resection success rate was 96.9% (95% confidence interval [CI] = 96.4% to 97.4%), and the overall false-negative rate was 9.5% (95% CI = 7.4% to 12.0%), with no statistical differences between training methods. Overall audit outcomes were excellent in all four categories. For all three training groups combined, a statistically significant positive association was observed between surgeons' average number of procedural errors and their false-negative rate ($\rho = +0.188$, $P = .021$).

Conclusions All three training methods resulted in uniform and high overall sentinel lymph node resection rates. Subgroup analyses identified some variation in false-negative rates that were related to audited outcome performance measures.

J Natl Cancer Inst 2009;101:1356–1362

Substantial side effects are associated with axillary lymph node resection, a surgical procedure that was originally designed to maximize breast cancer survival, provide regional control, and determine the stage of the patient's cancer. Sentinel lymph node surgery may offer equivalent outcomes with decreased side effects. The National Surgical Adjuvant Breast and Bowel Project (NSABP) B-32 trial was designed to determine whether sentinel lymph node resection achieves the same therapeutic outcomes as axillary lymph node resection but with fewer side effects. The primary endpoints of the B-32 trial are survival, regional control, and morbidity. Patients are currently being monitored for these outcomes, and definitive analysis is not yet available. Secondary outcomes related to the technical issues of sentinel lymph node resection and patient characteristics have recently been reported (1).

In a validation study that preceded the B-32 trial (2), a potential for variability was observed as to how the sentinel lymph node procedure was performed and how source documentation related to the sentinel lymph node procedure was prepared. In contrast to systemic

Affiliations of authors: College of Medicine, University of Vermont, Burlington, VT (DNK, TA, SPH, JMS, DLW); National Surgical Adjuvant Breast and Bowel Project, Pittsburgh, PA (TBJ, NW); Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA (AMB); Department of Human Oncology, Allegheny General Hospital, Pittsburgh, PA (TBJ, NW).

Correspondence to: David N. Krag, MD, Department of Surgery, College of Medicine, University of Vermont, Given Building, Rm E309, 89 Beaumont Ave, Burlington, VT 05405 (e-mail: david.krag@uvm.edu).

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djp281

Advance Access publication on August 24, 2009.

therapy trials in which administration of drugs is well documented, surgical procedures are not necessarily documented in a standardized manner. Source documentation relevant to a surgical trial may be incomplete. This documentation may be more challenging when the procedure is new and of high technical complexity. To minimize procedural variation in the B-32 trial, standardization of methods and of auditing was instituted for sentinel lymph node resection, pathological processing, generation of source documents, and data entry.

Given the large target accrual for the B-32 trial, the goal was to train a sufficiently large number of surgeons to account for possible attrition. A plan was instituted in which site visits, which were based on the methods used in the validation study, were performed in a standardized manner (3). A set of core instructors were trained to perform site visits by the principal investigator of this protocol (D. N. Krag) and by the Surgical Training Chair (S. P. Harlow). The site visit included intraoperative educational review by a designated surgeon. After a surgeon had been approved to enroll patients in the B-32 trial, that surgeon, now termed a site trainer, could then perform the intraoperative evaluation of other surgeons at that site.

Surgeons were required to demonstrate compliance with the surgical protocol, generation of source documentation, and accuracy of data entry. The time points for demonstrating compliance were before randomization, after completion of the first 10 surgeries, and after completion of the second 10 surgeries; these data have been reported previously (3). It was also demonstrated previously (2) that successful completion of five surgeries according to protocol guidelines was sufficient for a group of surgeons with limited experience to achieve an overall group rate for sentinel lymph node resection of approximately 90%. This number of surgeries was initially required for the B-32 trial. However, over the course of the B-32 trial, surgeons who had extensive experience in sentinel lymph node surgery registered for participation. An expedited approach was, thus, instituted for these surgeons. The training method was unchanged, but documentation of successful protocol compliance required one rather than five successful surgeries before the surgeon could accrue patients to the B-32 trial.

The purpose of this report was to assess the effectiveness of the training methods, overall protocol compliance, and their relationship to technical outcomes. Outcomes were further analyzed according to the required number of operations (one vs five operations) before a surgeon was approved to randomly assign patients to treatment groups. The group of surgeons with five operations before randomization was further subdivided into two groups according to intraoperative evaluation by a core instructor or by a site trainer.

CONTEXT AND CAVEATS

Prior knowledge

The randomized National Surgical Adjuvant Breast and Bowel Project B-32 trial is evaluating whether sentinel lymph node resection can achieve the same outcomes as axillary lymph node resection but with fewer side effects.

Study design

The 'overall' relationship of surgeon trial preparation, protocol compliance audit, and technical outcomes was 'determined and' compared among surgeons who were trained in one of the three ways. Surgical performance was audited in four categories.

Contribution

No statistically significant differences were observed between training methods: Overall audit outcomes were excellent in all four categories. Among all surgeons, a statistically significant positive association was observed between the average number of procedural errors and the false-negative rate. Some variation in false-negative rates was observed that was related to audited outcomes.

Implications

Training methods were effective. Variation in false-negative rates in subgroup analyses indicates the value of the auditing measures and supports the use of similar auditing measures in future trials.

Limitations

Before randomization, there was only one intraoperative educational session per surgeon to ensure awareness of all the steps involved in protocol compliance. Audits of randomized cases were limited to 20 operations.

From the Editors

Participants and Methods

The NSABP trial B-32 (ClinicalTrials.gov., NCT00003830) was undertaken after approval from local institutional review boards and in accord with an assurance filed with and approved by the US Department of Health and Human Services. Informed written consent was obtained from each participant in this study.

Summary of NSABP B-32 Trial Design

Patients with operable invasive breast cancer and clinically negative axillary lymph nodes were randomly assigned to receive either sentinel lymph node resection followed by immediate conventional axillary lymph node resection of the remaining nonsentinel lymph nodes (group 1) or sentinel lymph node resection without axillary lymph node resection if sentinel lymph nodes were negative on intraoperative cytology and histological examination (group 2; Figure 1). Patients in group 2 underwent axillary lymph node resection only if no sentinel lymph nodes were identified or if one or more sentinel lymph nodes were positive on intraoperative cytology or subsequent histological examination.

The primary endpoints of the B-32 trial are survival, regional control, and morbidity. The number of events for the definitive analyses of these endpoints has not yet been reached, and patients are being carefully monitored. Two secondary endpoints, accuracy and technical success, are complete and have been reported previously (1).

From May 1, 1999, through February 29, 2004, 5611 patients were entered and randomly assigned to a treatment group by 233 surgeons from 80 institutions in the United States and Canada. A total of 224 surgeons received one of the three types of training and certification by one of the nine B-32 core study trainers. These 224 trained surgeons, then, enrolled a total of 4994 patients (of total 5611 patients) into the B-32 trial and are the subject of this report.

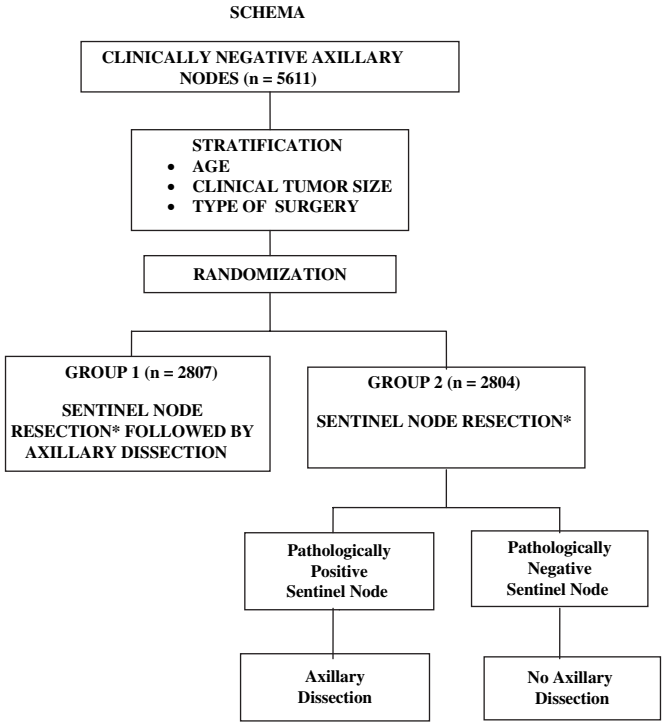

Figure 1. Randomization schema of the B-32 trial. Asterisk indicates that in this group, patients in whom a sentinel lymph node was not identified received an axillary lymph node dissection.

Training Methods

The key elements of training included 1) a detailed manual that included the study protocol design, methods for sentinel lymph node surgery, instructions for generation of source documents and filling out data forms, and methods of labeling sentinel lymph node specimens; 2) a site visit by a core instructor that included review of the protocol with the key participants, including the surgeons, nuclear medicine physician, pathologist, and operating room nursing personnel; and 3) intraoperative educational session with the designated surgeon. Subsequent certification for approval to enroll patients in the B-32 trial was based on detailed evaluation of a minimum set of operations for procedural compliance, generation of source documentation, and data entry.

Core trainers were accompanied on their first site visit by the NSABP protocol principal investigator or by the Surgical Training Chair. The purpose was to standardize the site visit process. During a site visit by a core trainer, at least one surgeon had an intraoperative educational session. During this session, an explicit set of steps related to identification of the sentinel lymph node was reviewed. Also reviewed were the sentinel lymph node pathology labeling methods and the specific information to include in the operative note. Surgeons subsequently submitted documentation that was audited for protocol compliance. Surgeons (categories 1 and 2) who successfully completed review of submitted material from five operations before randomization were approved to accrue patients and randomly assign patients to a treatment group (3). Surgeons with more extensive experience (category 3) had an expedited review and were approved to accrue and randomly assign patients after successful completion and review of submitted materials from one operation.

Once approved to randomly assign patients to a treatment group, the surgeon was then eligible to act as a site trainer and conduct intraoperative education of other surgeons at that site. The newly trained surgeons at that site then submitted documentation of the appropriate number of surgeries before randomization that was audited for protocol compliance in the same manner as the other trained surgeons.

In this study, surgeons were first categorized according to the number of operations they had to perform before randomization: The expedited group was required to perform one operation and the standard group was required to perform five operations. Surgeons in the standard training group were secondarily categorized according to intraoperative evaluation (core instructor vs site trainer).

Surgical Procedure

Both technetium-99m sulfur colloid and isosulfan blue were used as described previously (1). The radioactive tracer was injected into the breast around the tumor from 30 minutes to 8 hours before surgery. Blue dye was injected into the breast around the tumor 5 minutes before incision. A gamma probe and visual guidance of blue-stained ducts were used to surgically identify sentinel lymph nodes. Lymphoscintigraphy was not required for this trial. Intraoperative evaluation of sentinel lymph nodes for patients in group 2 was performed first by cytology and subsequently for permanent analysis with hematoxylin–eosin staining. Immunohistochemistry was reserved for evaluation of cells considered questionable on hematoxylin–eosin slides. A complete axillary lymph node resection was performed if metastases were found by cytology or on hematoxylin–eosin slides.

Description of Data Forms and Source Documentation

Source documentation was defined as data available in the medical record. These data were entered in the operative note by the surgeon and into the pathology report by the pathologist. The training manual had a checklist of key data items to be entered into source documentation. A form for the collection of surgical pathology data was submitted to the NSABP, and this form was reviewed for accuracy on the basis of source documentation.

Audit Procedures

Audit criteria included 94 specific items in the following four categories: procedural, operative note, pathology report, and data entry (Table 1). Twenty-five specific items were assessed in relation to protocol compliance of the surgical procedure. Source documentation was audited by evaluating 14 items in the operative note and 11 in the pathology report. The data entry form had 44 fields that were scored for accuracy. When surgeons had performed 10 and 20 operations, continued permission to accrue and randomly assign patients to a treatment group was contingent upon successful completion of the audit. Audits were performed twice. When a surgeon successfully completed two audits, that surgeon was not audited further.

Definition of Scoring Methods for Audit-Related Outcomes

Any errors or omissions found in a screened field were recorded as an unweighted error occurring in that field. The fields were scored as having an error if the available source documentation did not include information that supported the information in the data form field.

Statistical Methods

The 95% confidence intervals (CIs) for sentinel lymph node resection rates and false-negative rates were calculated with the Clopper–Pearson approach (4). Patient-level pooled comparisons for surgeon groups used exact contingency table methods (StatXact, version 4; Cytel Software Corporation, Cambridge, MA). Audit outcomes per patient for each of the four categories were summarized overall as means and 95% confidence intervals. Spearman rank correlation coefficients (ρ) were used to examine the association of sentinel lymph node resection rates and false-negative rates to audit performance for each of the four audit categories. Surgeon-level audit performances were classified into quartiles. For each quartile, the overall average false-negative rates and the overall average number of sentinel lymph nodes removed are presented in Table 2. The contribution of individual procedural error items to the overall procedural error performance score for surgeons was examined by using a Cochran–Armitage linear trend test for 2 × 4 contingency tables, in which the row classification represented the individual item and the columns represented the surgeon performance quartiles. Linear or generalized linear mixed models were used to compare the three surgeon training groups relative to the average number of audit errors, sentinel lymph node resection rates, false-negative rates, and number of sentinel lymph nodes removed (because surgeons were classified under training group and patients were nested under surgeon). StatXact (version 4), SAS (version 9.1; SAS Institute, Inc, Cary, NC), and BMDP5V (BMDP Biomedical Computer Programs P-Series, University of California Press, Berkeley, CA) were used for model development, along with a 5% statistical significance level for formal testing. All statistical tests were two-sided.

Results

Overall Results

A total of 572 surgeons registered for training and 369 of them were trained. The number approved to randomly assign patients to surgery group was 261, and 224 of these 261 surgeons enrolled 4994 of the 5611 patients in the B-32 trial (group 1 = 2483 patients and group 2 = 2511 patients). Among the 224 surgeons, 121 were core trained, 50 were site trained, and 53 received expedited training (Table 3).

The 224 trained surgeons had an overall sentinel lymph node success rate of 96.9% (95% CI = 96.4% to 97.4%), with no statistically significant differences among the three training groups. False-negative rates for 153 surgeons were calculated on the basis of the 682 patients in group 1 who had at least one lymph node (sentinel lymph node or nonsentinel lymph node) that was pathologically positive for metastases and had an axillary lymph node dissection. The overall false-negative rate was 9.5% (95% CI = 7.4% to 12.0%), and no statistically significant differences in the false-negative rates were observed between the three training groups (P = .142).

Audit Outcomes: Procedural Errors, Operative Notes, Pathology Reports, and Data Entry

Data from a total of 2493 operations from 217 of the 224 surgeons were audited. The seven surgeons who were not audited had each accrued fewer than 10 patients to the trial. The overall mean number of errors and 95% confidence intervals identified per operation for each of the four audit categories were as follows: procedural 0.64 (95% CI = 0.60 to 0.67), operative note 0.50 (95% CI = 0.46 to 0.54), pathology report 0.54 (95% CI = 0.50 to 0.58), and data entry 1.62 (95% CI = 1.54 to 1.70). The four audited categories for the three training groups were uniformly good, with procedural compliance, operative notes, and pathology report errors averaging well below one error per patient. Data entry errors averaged between one and two errors per patient.

Success Rate of Sentinel Lymph Node Resection and Audit Outcomes

The success rate of sentinel lymph node resection per surgeon was evaluated for all patients according to surgeon training and audited performance outcomes to determine if any of the four audited outcomes was related to sentinel lymph node resection rates. Sentinel lymph node resection rates were high, and the audit outcomes were good (as described above). No correlation between variations in success rates according to audit outcomes was observed.

False-Negative Rate and Audit Outcomes

For all three training groups combined, a statistically significant positive association was observed between surgeons' average number of procedural errors and their false-negative rate (ρ = +0.188, P = .021) (Table 2). Of the 25 audited procedural variables, nine had statistically significant linear trends in audit performance (P < .001) (Table 4).

False-Negative Rate and Number of Sentinel Lymph Nodes Removed

False-negative rates were determined for the 153 surgeons (by use of the 682 patients in group 1 with at least one positive lymph node) according to training type and number of lymph nodes removed. Among patients with one sentinel lymph node removed, the overall false-negative rate was 18.5%. The rate then declined to 4.4% among patients with four or more sentinel lymph nodes removed (Table 5). By use of a generalized linear mixed model analysis, a statistically significant decrease in the false-negative rate was observed, with increasing numbers of sentinel lymph nodes removed (P = .010) for all training groups.

Number of Sentinel Lymph Nodes Removed and Training Group

The number of sentinel lymph nodes removed by each training group of surgeons from 4839 of the 4841 patients was determined; the number of sentinel lymph nodes removed was not recorded for two patients. There was a statistically significant difference in the overall number of sentinel lymph nodes removed among the three surgeon training groups (P < .001). Surgeons in the expedited training group removed the largest number of sentinel lymph nodes, 3.22 (95% CI = 3.07 to 3.37), followed by the core-trained surgeons who removed 2.84 (95% CI = 2.77 to 2.90) lymph nodes, and then the site-trained surgeons who removed 2.60 (95% CI = 2.45 to 2.76) lymph nodes.

Table 1. Audit criteria data categories*

Procedural events	Operative notes	Pathology reports	Data entry
Patient eligibility (1)	Location of tumor in breast (1)	Specimen identifiers (1)	Primary breast surgery information (7)
TSC injection procedure (4)	Hot spot information (6)	Location, hot spots, ex vivo counts (3)	TSC injection information (2)
Blue dye procedure (3)	Timing of TSC injection (1)	Presence of blue dye (1)	Hot spot identification and bed count (7)
Saline injection procedure (3)	Volume of dye injection (1)	Axillary dissection identification (1)	SN identifiers (8)
Axillary LN dissection level and appropriateness (3)	Presence of blue dye in LN or duct (1)	Histological grade (1)	SN and axillary LN dissection surgery information (4)
Hot spot exploration (3)	Ex vivo counts of SNs (1)	Receptor status (1)	SN intraoperative interpretation (2)
Bed count readings and procedure (4)	Location of SNs (1)	Tumor type (1)	SN and axillary LN dissection final diagnosis (7)
SN identified, No. of LNs (2)	Intraoperative cytology (1)	Maximum pathological tumor size (1)	Breast tumor information (5)
False-negative and other findings (2)	Axillary LN dissection performed (1)	Results of intraoperative cytology (1)	Adverse events (2)

* For procedural events, 25 specific items were assessed in relation to protocol compliance of the surgical procedure. For operative notes and pathology reports, source documentation was audited by evaluating 14 items in the operative note and 11 in the pathology report. For data entry, 44 fields were scored for accuracy. Numbers in parentheses indicate the number of data fields for each described category (total data fields = 94). LN = lymph node; SN = sentinel lymph node; TSC = Technetium-99m sulfur colloid.

Number of Sentinel Lymph Nodes Removed and Audited Outcomes

For all three training groups combined, statistically significant negative associations were observed between the surgeons' average number of procedural errors and the number of sentinel lymph nodes removed ($\rho = -0.254$, $P < .001$). The negative relationship of the procedural audit measures and number of lymph nodes removed is shown in Table 2.

Discussion

The most important results of this study are related to surgeon performance and to trial participation. Surgeon performance was measured by overall success rate of sentinel lymph node resection, which was high at 96.9% and the overall false-negative rate at 9.5%. There were no statistical differences between the different training groups. The audited outcomes for all training groups were excellent and with procedural, operative note, and pathology report errors averaging below one error per patient and data entry errors between one and two errors per patient. Subgroup analysis of surgeon performance identified an association between the false-negative rate and the average number of procedural errors for all three groups combined. There was also an association between the number of sentinel lymph nodes removed and the procedural errors. The higher the number of procedural errors, the lower the number of sentinel lymph nodes removed. Trial participation was reflected in the decreasing number of surgeons that signed up to participate, were trained, and enrolled patients in the trial.

The CONSORT statement was developed to improve the design and conduct of randomized clinical trials (5). There are, however, unique issues related to surgical trials that are not addressed by the CONSORT recommendations. For example, variation in performance of a surgical procedure and documentation of procedures explicitly related to the surgical aspect of the trial are not addressed, yet are important issues. When the B-32 trial was launched, there were no established universal standards for minimizing variation in surgical performance. Data obtained in a sentinel lymph node study (2) preceding the B-32 trial indicated that there was potential for variation in surgical performance. It was also previously observed that there were unique challenges of documenting surgery-related data (2). The B-32 trial was well controlled for patient variables (1), and a program was designed and implemented for controlling surgeon variables. Surgeons participated in on-site education and auditing before and after enrollment. A checklist of important steps in the surgical procedure was used and was also a guide for incorporating elements into source documentation. Audit outcomes were designed to assess the following issues: procedural compliance with the sentinel lymph node resection protocol, generation of source documentation both surgical and pathological, and data entry. These considerable efforts were made to ensure that interpretation of the outcomes observed between the two randomized arms would be as free as possible from variation. This is particularly

Table 2. Audit performance, the false-negative rate, and the number of sentinel lymph nodes removed*

	Quartiles of procedural errors			
Variable	**1**	**2**	**3**	**4**
False-negative rate, mean (95% CI)	4.3 (0.7 to 7.9)	5.8 (1.5 to 10.0)	20.7 (9.5 to 31.8)	16.0 (6.2 to 25.8)
No. of SNs removed, mean (95% CI)	2.97 (2.67 to 3.26)	2.98 (2.68 to 3.27)	2.67 (2.39 to 2.95)	2.37 (2.14 to 2.61)

* CI = confidence interval; SN = sentinel lymph node.

Table 3. Sentinel lymph node resection success rates and false-negative rates for 224 trained surgeons according to training method*

	Technical success†				False-negative results‡			
Type of training	**No. of surgeons**	**n**	**N**	**Rate, % (95% CI)**	**No. of surgeons**	**n**	**N**	**Rate, % (95% CI)**
Core trained	121	3411	3527	96.7 (96.1 to 97.3)	90	52	493	10.6 (8.0 to 13.6)
Site trained	50	514	525	97.9 (96.3 to 99.0)	31	7	64	10.9 (4.5 to 21.3)
Expedited	53	916	942	97.2 (96.0 to 98.2)	32	6	125	4.8 (1.8 to 10.2)
Overall	224	4841	4994	96.9 (96.4 to 97.4)	153	65	682	9.5 (7.4 to 12.0)

* CI = confidence interval.

† For technical success, n = number of operations in which sentinel lymph nodes were successfully resected; N = total number of operations for that category. $\chi^2 = 2.553$; $df = 2$; $P = .286$. All statistical tests were two-sided.

‡ For false-negative results, n = number of operations in which sentinel lymph nodes were negative but nonsentinel lymph nodes were pathologically positive; N = total number of operations in that category with any pathologically positive lymph node. $\chi^2 = 3.983$; $df = 2$; $P = .142$. All statistical tests were two-sided.

important because the differences expected in primary outcomes between the two groups may be very small.

The overall success rate of sentinel lymph node resection was high at 96.9%, and there were no statistical differences observed between trained groups. These results indicate that the B-32 training program allowed surgeons from the three trained groups to achieve an overall equal and high success rate throughout the trial. Surgeons were no longer audited after successful completion of two audits, and the sustained high sentinel lymph node resection rates indicate that compliance did not materially decrease after the audit period.

The overall false-negative rate was 9.5%, and there was no statistical difference in the false-negative rates between the different training groups. These results indicate that the training program resulted in surgeons achieving an equivalent overall false-negative rate across all three training methods. This rate is slightly lower than the 11% false-negative rate reported in the initial validation study preceding the B-32 trial (2) but is also consistent with the average false-negative rate of 8.4% (range = 0% to 29%) that has been reported in a meta-analysis of 69 studies (6).

Limitations of this study are related to duration of audit and direct objective observation of surgeries. Ideally, all surgeries would be audited, but application of the level of auditing performed in this study to all surgeries is impractical and expensive. After 20 surgeries, it is unlikely that audited surgeons would perform with substantial differences during later surgeries. Intraoperative observation of all surgeries provides objective documentation of protocol performance, but this is not feasible from a resource perspective. Our approach with a single intraoperative observation was to ensure that the surgeon was aware of all the steps involved to perform sentinel lymph node surgery according to protocol and importantly could document performance of the steps, which provide written validation that the protocol was followed. For example, if the surgeon documented that the bed count was low after removal of a radiolabeled sentinel lymph node, this act was good assurance that the full extent of sentinel lymph nodes had been removed. The study design included several critical steps to be documented by the surgeon for all surgeries. The quality of this approach is validated by identification of a relationship of higher false-negative rates to lower audit performance in a subset of surgeons.

The audited outcomes for the three training groups demonstrated that the procedural, operative note, and pathology report errors averaged well below one error per patient. Data entry errors averaged between one and two errors per patient. Given that there were 94 audited data elements evaluated for errors, the outcomes were excellent and this level of protocol compliance is infrequently matched in the surgical literature (7–9).

Several variables were examined to determine whether there was any subgroup variation of the false-negative rate relative to audit outcomes. An association was observed between the false-negative rate and the average number of procedural errors for all three groups combined. Nine of the 25 individual audit measures

Table 4. Procedural audit measures that reflect a linear trend across surgeon performance quartiles*

	Quartile of procedural errors			
Procedural audit measure	**1 (n = 575)**	**2 (n = 578)**	**3 (n = 543)**	**4 (n = 508)**
TSC injected on four sides of the tumor, No. (%)	7 (1.2)	8 (1.4)	22 (4.1)	25 (4.9)
Volume of TSC, No. (%)	11 (1.9)	39 (6.7)	53 (9.8)	58 (11.4)
Dose TSC, No. (%)	5 (0.9)	25 (4.3)	23 (4.2)	55 (10.8)
Volume of blue dye, No. (%)	1 (0.2)	7 (1.2)	9 (1.7)	35 (6.9)
Blue dye injection before survey, No. (%)	2 (0.43)	5 (0.9)	29 (5.3)	112 (22.0)
Saline used if necessary, No. (%)	0 (0.0)	2 (0.43)	7 (1.3)	26 (5.1)
Bed count >10% of hottest node or no source document, No. (%)	3 (0.5)	15 (2.6)	23 (4.2)	24 (4.7)
No hot spot identified, No. (%)	15 (2.6)	37 (6.4)	58 (10.7)	72 (14.2)
Other, No. (%)	9 (1.6)	43 (7.4)	83 (15.3)	139 (27.4)

* For the 152 audited surgeons who had at least one patient with a positive lymph node, nine of the 25 individual audit measures had a Cochran–Armitage trend test *P* value of less than .001 that reflected the linear trend across surgeon performance quartiles. These nine audit measures are listed in this table. Percentage of operations with procedural error is listed separately for each quartile on the basis of the mean number of procedural errors. n = the number of operations; TSC = Technetium-99m sulfur colloid.

Table 5. False-negative rate for 153 surgeons from all training groups (on the basis of 682 patients in group 1 with at least one positive lymph node) according to the number of sentinel lymph nodes removed*

No. of SNs removed	No. of false negative/ total	False-negative rate, % (95% CI)
1	28/151	18.5 (12.7 to 25.7)
2	17/157	10.8 (6.4 to 16.8)
3	11/168	6.6 (3.3 to 11.4)
≥4	9/206	4.4 (2.0 to 8.1)

* CI = confidence interval; SN = sentinel lymph node; total = number of patients with positive nodes according to the number of SNs removed.

reflected the linear trend across surgeon performance quartiles. These variables are related to technical aspects of the procedure that could potentially affect false-negative rate. For example, only 2.6% of the operations performed by surgeons in the best quartile for procedure-related audit were unable to find a preincision hot spot compared with 14% in the lowest quartile. In general, this would indicate that fewer procedural errors tended to result in lower false-negative rates. These data would support the development of a standardized procedural checklist that is followed sequentially during surgery and could serve as a performance reference for documentation purposes.

We also found an association between the number of sentinel lymph nodes removed and the procedural errors. The higher the number of procedural errors, the lower the number of sentinel lymph nodes removed. The number of sentinel lymph nodes removed has been reported previously to be a variable that is strongly and statistically significantly associated with the false-negative rate (10,11). Thus, surgeons should make a diligent effort to remove all sentinel lymph nodes, and the number of sentinel lymph nodes removed should be considered as a variable for documenting procedural compliance.

We also found statistically significant differences among the groups of surgeons with respect to the number of sentinel lymph nodes removed. Surgeons in the expedited group removed more sentinel lymph nodes than those in either of the two other trained groups. This result may contribute to the apparent difference (although not statistically significantly so) in false-negative rates between the trained groups. There were not enough false-negative data to attribute variation with statistical significance.

Interest in participating in the B-32 trial was high. A total of 572 surgeons registered to participate through NSABP trial sites, but only 369 registered surgeons were trained, despite the availability of trainers. Among the group that was trained, only 224 surgeons enrolled patients. It appears that the complexity of participating in the trial and the lack of experience in enrolling patients in randomized trials may have, at least partially, affected surgeon participation. These observations are relevant to feasibility and resource assessment of future randomized surgical trials.

In conclusion, the B-32 trial represents one of the most carefully controlled and monitored randomized trials in the field of surgical oncology. The training methods used for the B-32 trial were effective and resulted in uniform and high overall sentinel lymph node resection rates and false-negative rates. Overall, no statistically significant differences were observed in sentinel lymph node resection rates and false-negative rates between the three training groups of surgeons. Subgroup analysis identified some variation in false-negative rates that were related to audited outcome performance measures, indicating the value of similar auditing measures on future trials.

References

1. Krag DN, Anderson SJ, Julian TB, et al. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. *Lancet Oncol.* 2007;8(10):881–888.
2. Krag D, Weaver D, Ashikaga T, et al. The sentinel node in breast cancer—a multicenter validation study. *N Engl J Med.* 1998;339(14):941–946.
3. Harlow SP, Krag DN, Julian TB, et al. Prerandomization Surgical Training for the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-32 trial: a randomized phase III clinical trial to compare sentinel node resection to conventional axillary dissection in clinically node-negative breast cancer. *Ann Surg.* 2005;241(1):48–54.
4. Conover WJ. *Practical Nonparametric Statistic.* New York, NY: John Wiley & Sons; 1971.
5. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA.* 1996;276(8):637–639.
6. Kim T, Giuliano AE, Lyman GH. Lymphatic mapping and sentinel lymph node biopsy in early-stage breast carcinoma: a metaanalysis. *Cancer.* 2006;106(1):4–16.
7. You YN, Jacobs L, Martinez ED, et al. Improved surgeon performance in clinical trials: an analysis of quality assurance audits from the American College of Surgeons Oncology Group. *J Am Coll Surg.* 2006;203(3):269–276.
8. Sasako M. Clinical trials of surgical treatment of malignant diseases. *Int J Clin Oncol.* 2005;10(3):165–170.
9. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg.* 2006;244(5):663–667.
10. Povoski SP, Young DC, Walker MJ, et al. Re-emphasizing the concept of adequacy of intraoperative assessment of the axillary sentinel lymph nodes for identifying nodal positivity during breast cancer surgery. *World J Surg Oncol.* 2007;5:18.
11. Wong SL, Edwards MJ, Chao C, et al. Sentinel lymph node biopsy for breast cancer: impact of the number of sentinel nodes removed on the false-negative rate. *J Am Coll Surg.* 2001;192(6):689–691.

Funding

National Cancer Institute, Department of Health and Human Services, Public Health Service Grants (P30 CA022435 Vermont Cancer Center, 5RO1CA74137-9 D.N.K., U10-CA-12027, U10-CA-69651, U10-CA-37377, and U10-CA-69974 to National Surgical Adjuvant Breast and Bowel Project).

Notes

The authors have full responsibility for design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript.

We thank Elaine Elliott, Barbara Branch, and Sarah Howe for assistance in performing this study.

Manuscript received June 25, 2009; revised June 25, 2009; accepted July 21, 2009.