Quality Measures for Breast Cancer Surgery

Initial Validation of Feasibility and Assessment of Variation Among Surgeons

Laurence E. McCahill, MD; Alicia Privette, MD; Ted James, MD; Johanna Sheehey-Jones, RN; John Ratliff, JD; Donald Majercik, MD; David N. Krag, MD; Mary Stanley, MD; Seth Harlow, MD

Objectives: To identify and quantify surgical outcomes as possible quality measures of initial breast cancer surgery and to assess variation among surgeons.

Design: Descriptive analysis of concurrently collected outcome measures.

Setting: University hospital with a designated breast cancer center.


Main Outcome Measures: Eight measures were identified: (1) total mastectomy rate; (2) close (<1 mm) and positive margin rate following initial partial mastectomy; (3) number of operations required in breast conservation; (4) number of nodes obtained from sentinel lymph node biopsy; (5) number of nodes from axillary dissection; (6) proportion of patients with positive sentinel lymph node biopsy undergoing axillary dissection; (7) use of intraoperative lymph node assessment; and (8) time from diagnosis to surgery.

Results: Nine hundred ten operations (218 for ductal carcinoma in situ, 692 for invasive breast cancer) were performed by 6 surgeons. Variation existed among surgeons in the combined close and positive margin rate, number of nodes obtained from sentinel lymph node biopsy, and use of intraoperative lymph node assessment. No significant variation was seen for the overall mastectomy rate, mean number of operations, positive margin rate alone, and number of lymph nodes from axillary dissection.

Conclusions: Quality indicators for breast cancer surgery can be identified and readily monitored. Outcome variation exists at a high-volume breast center. Further study into the causes and effects of this variation on short- and long-term patient outcomes as well as health care costs is needed.

Arch Surg. 2009;144(5):455-462

Breast cancer remains the most common malignant neoplasm facing the general, oncologic, or dedicated breast surgeon. Breast cancer is a high-incidence disease, with 185,000 patients diagnosed annually in the United States, making breast cancer surgery the most common cancer operation in the United States (> 200,000 operations/y).

Breast cancer surgery, collectively referring to surgical management of the primary tumor and axilla, remains a mainstay of breast cancer treatment. Patients may routinely undergo multiple procedures in short sequence for their primary tumor to obtain clear margins as well as axillary surgery for removal of lymph nodes as either a staging procedure or a potentially therapeutic intervention. Variability in mastectomy rates has been previously demonstrated. Variation in surgical performance of partial mastectomy or sentinel lymph node biopsy (SLNB) has been less frequently evaluated. How often and to what degree patients with breast cancer undergo multiple operations are not clear from the literature but may be useful as quality measures. Patients with newly diagnosed breast cancer may be better informed if more complete and detailed surgical outcomes are available at the time of their treatment decision making.

Quality assessment for lower-risk operations is particularly complex as easily measurable outcomes (ie, death, major morbidity) are unlikely to vary for low-risk procedures. A good proposed quality measure must be readily measurable, be clinically meaningful, and have likely variation among clinicians or hospitals.

See Invited Critique at end of article

Author Affiliations: Division of Surgical Oncology, Department of Surgery (Drs McCahill, Privette, James, Majercik, Krag, Stanley, and Harlow) and James M. Jeffords Institute for Quality (Ms Sheehey-Jones and Dr Ratliff), University of Vermont and Fletcher Allen Health Care, Burlington.
Few guidelines exist regarding the performance and outcomes of breast cancer surgery. These guidelines have not been widely used for several reasons, such as controversy of what constitutes an adequate pathologic margin following partial mastectomy. Furthermore, standard health care administrative databases do not accurately capture multiple clinical factors (i.e., tumor size, patient treatment preference, patient family history) that most surgeons feel contribute significantly to the choice of initial breast cancer surgery. The lack of well-accepted guidelines or any standardized reporting of breast cancer surgery outcomes may result in patients receiving widely variable surgical treatment based on geographic location, choice of hospital, and surgeon.

To propose meaningful quality measures, a quality-monitoring program was initiated to evaluate initial breast cancer surgery performed at Fletcher Allen Health Care. The purposes of this study were to establish the feasibility of quality assessment to determine ranges of outcomes among surgeons and to quantitatively examine variation at our institution. By defining specific outcome measures and quantifying variation among surgeons, we hoped to identify opportunities for collaboration and improvements in surgical technique or breast cancer management strategies at our institution. We also hoped to provide realistic measures that can be adopted nationally to gain a better understanding of quality and value in health care as it applies to current breast cancer surgical care.

**METHODS**

Following institutional review board approval by the University of Vermont, we identified patients using a quality database for patients who underwent initial breast cancer surgical treatment at our institution from April 1, 2003, to March 30, 2008. Data were concurrently collected on numerous clinical parameters and specific short-term surgical outcomes for all of the surgeons performing breast cancer surgery (L.E.M., T.J., D.M., D.N.K., M.S., and S.H.).

Inclusion criteria were patients with a preoperative diagnosis of breast cancer (ductal carcinoma in situ or invasive ductal or lobular cancer) receiving their initial cancer operation at our institution. Patients with a history of breast cancer treatment longer than 1 year previously were included. Exclusion criteria included patients with a preoperative diagnosis of lobular cancer in situ or atypical ductal hyperplasia. Patients who underwent initial breast surgery at an outside institution and required subsequent surgery were excluded. Nurse breast health specialists identified eligible patients and collected data prospectively using a standardized data entry form after training and with monitoring by one of us (L.E.M.). Data were entered into an electronic database.

We prospectively identified 8 outcome measures at the onset of this project that might reflect on the quality of breast cancer surgery. Three related to management of the primary tumor, 4 related to management of the axilla, and 1 related to efficiency of health care delivery. Measures related to primary tumor management were as follows: (1) total mastectomy (TM) rate; (2) close (<1 mm) and positive margin rate following initial partial mastectomy (PM); and (3) mean number of operations performed per patient to complete primary surgical management for cancer. Measures related to management of the axilla were as follows: (4) number of lymph nodes obtained during SLNB; (5) number of nodes obtained in axillary dissection (AD); (6) use of intraoperative assessment for sentinel nodes; and (7) proportion of patients with positive SLNB undergoing AD. Efficiency of surgical cancer care was measured as follows: (8) days from pathologic diagnosis to initial breast cancer surgery (excluding patients undergoing neoadjuvant chemotherapy for this measure).

The initial TM rate was defined as TMs performed as a first procedure divided by the initial breast cancer surgical procedures performed. To control for variation in patient medical history, an adjusted TM rate excluded patients ineligible for breast conservation (a history of irradiation, multicentric breast cancer, etc.). The final adjusted TM rate included initial TMs as well as any TMs following initial attempts at breast conservation. The TMs performed on a contralateral breast for prophylaxis were excluded.

The combined close and positive margin rate was defined as the number of initial PMs performed with a close (tumor <1 mm from any inked margin) or positive (tumor at inked margin) margin on final pathologic examination divided by the number of initial PMs performed. The vast majority of primary breast malignant pathologic results are reviewed by the Fletcher Allen Health Care breast pathology team and re-reviewed in a breast cancer multidisciplinary conference. Breast nurse specialists were trained in the interpretation and recording of pathologic margins. Mean operations performed for attempted breast conservation were the number of operations required to complete definitive surgical management and included initial PMs and subsequent reexcisions as well as any subsequent TMs.

For outcomes related to axillary surgery, patients with prior ipsilateral breast cancer were excluded as most had prior AD, which would confound the results. A positive SLNB was defined as the identification of either micrometastases (0.2-2.0 mm) or metastases greater than 2 mm on final pathologic examination. Isolated tumor cells (<0.2 mm) were recorded as node negative. A sentinel lymph node initially identified as positive on intraoperative touch preparation cytology but negative on final pathologic examination was considered negative. A completion AD could be performed at the time of initial breast surgery or as a secondary procedure.

Outcomes were compared between individual surgeons for each of the 8 measures. Surgeons were also divided into low-volume (<30 operations/y) and high-volume (≥30 operations/y) surgeons. Intrasurgeon variation within the 2 groups (low volume and high volume) was also assessed.

Data for each outcome measure were obtained using a parameter-specific database query. The data were then categorized by surgeon. Comparisons between surgeons for each parameter were calculated using χ² or Fischer exact analysis or nonparametric (Kruskal-Wallis) analysis as indicated. All of the calculations were completed using Stata version 8 statistical software (Stata Corp, College Station, Texas).

**RESULTS**

From April 1, 2003, until March 30, 2008, 910 eligible patients were identified. Approximately 90% of patients with breast cancer undergoing breast surgery at our institution were included. Most exclusions were for patients having initial surgery at an outside institution or for patients without confirmed malignant diagnosis preoperatively. The mean patient age was 58.4 years. The preoperative diagnosis was ductal carcinoma in situ for 218 patients (24.0%) and invasive cancer for 692 patients (76.0%). Final pathologic examination showed ductal carcinoma in situ alone in 181 patients (19.9%) and invasive cancer in 729 patients (80.1%), with 643 patients (88.2%) having invasive ductal cancer and 86 patients (11.8%) having invasive lobu-
lar cancer. The mean tumor size for all of the patients was 17.4 mm. The mean tumor size was 14.7 mm for patients undergoing initial PM and 25.3 mm for patients undergoing initial TM. Pathologic tumor size, histologic diagnosis, and presence of multicentric tumors did not vary among surgeons (Table 1).

Variability among surgeons for outcomes is shown in Table 2 and depicted graphically in the Figure. Four of the 6 surgeons received fellowship training in surgical oncology, and posttraining experience ranged from 3 to 25 years. Two lower-volume surgeons who had been in practice longer had a higher percentage of older patients and patients with previous ipsilateral breast cancer. For the initial operation, 235 patients (25.8%) underwent TM and 675 (74.2%) underwent PM. The initial TM rate and the adjusted initial TM rate did not vary among individual surgeons. The adjusted final TM rate and the adjusted initial TM rate did not vary among surgeons. The adjusted final TM rate was 24.4% (7.7%-34.2%) and the adjusted initial TM rate was 21.6% (7.7%-30.4%). Patients receiving primary systemic therapy prior to surgery were excluded. Patients undergoing completion AD following a positive SLNB had a broad range (47.4%-100.0%) but did not achieve statistical significance (P=.20). The mean number of nodes following AD was 16.9 and did not vary among surgeons.

Comparison of outcomes between high- and low-volume surgeons is shown in Table 4. The initial numbers of breast cancer operations performed annually by high- and low-volume surgeons were 57.4 and 7.9, respectively. Low-volume surgeons were noted to have an adjusted TM rate (both initial and final) and combined close and positive margin rate similar to those of high-volume surgeons. The only difference was a greater number of sentinel nodes identified by high-volume surgeons as compared with low-volume surgeons (3.8 vs 2.8, respectively; P<.001). Low-volume surgeons were more likely than high-volume surgeons to perform a completion AD following positive SLNB (84.6% vs 56.3%, respectively; P=.04).

### Table 1. Patient Clinical Factor Variance by Treating Surgeon

<table>
<thead>
<tr>
<th>Patient Demographic</th>
<th>Institutional Average</th>
<th>Surgeon Average, Range</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database-eligible operations/y, No.</td>
<td>182</td>
<td>3.6-70.6</td>
<td>NA</td>
</tr>
<tr>
<td>Patient age, y</td>
<td>58.4 (12.7)a</td>
<td>56.8-67.5</td>
<td>.001b</td>
</tr>
<tr>
<td>Final diagnosis, %</td>
<td></td>
<td></td>
<td>.08</td>
</tr>
<tr>
<td>Ductal carcinoma in situ</td>
<td>19.9</td>
<td>5.6-28.4</td>
<td></td>
</tr>
<tr>
<td>Invasive cancer</td>
<td>80.1</td>
<td>71.6-94.4</td>
<td></td>
</tr>
<tr>
<td>Final tumor size, mm</td>
<td>17.4 (17.6)a</td>
<td>13.9-19.2</td>
<td>.30</td>
</tr>
<tr>
<td>History of ipsilateral breast cancer, %</td>
<td>7.6</td>
<td>5.7-33.3</td>
<td>.001b</td>
</tr>
<tr>
<td>Presence of tumor in multiple quadrants, %</td>
<td>1.2</td>
<td>0.0-4.8</td>
<td>.10</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.

aData are expressed as mean (SD).

bNo significant variation if the 2 lowest-volume surgeons are removed (combined n=39).

### Table 2. Surgical Outcomes for Breast Cancer Surgery

<table>
<thead>
<tr>
<th>Quality Measure</th>
<th>Institutional Average</th>
<th>Surgeon Average, Range</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to surgery, d4</td>
<td>22.4</td>
<td>21.4-28.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Initial TM rate, % (n=910)</td>
<td>25.8</td>
<td>23.8-38.9</td>
<td>.15</td>
</tr>
<tr>
<td>Adjusted initial TM rate, % (n=827)b</td>
<td>21.6</td>
<td>7.7-30.4</td>
<td>.16</td>
</tr>
<tr>
<td>≤20 mm</td>
<td>15.4</td>
<td>0.0-23.9</td>
<td>.22</td>
</tr>
<tr>
<td>21-50 mm</td>
<td>32.4</td>
<td>23.8-64.3</td>
<td>.06</td>
</tr>
<tr>
<td>&gt;50 mm</td>
<td>67.4</td>
<td>60.0-100.0</td>
<td>.80</td>
</tr>
<tr>
<td>Adjusted final TM rate, % (n=827)b</td>
<td>24.4</td>
<td>7.7-34.2</td>
<td>.06</td>
</tr>
<tr>
<td>≤20 mm</td>
<td>17.2</td>
<td>0.0-23.9</td>
<td>.30</td>
</tr>
<tr>
<td>21-50 mm</td>
<td>35.8</td>
<td>25.4-71.4</td>
<td>.03</td>
</tr>
<tr>
<td>&gt;50 mm</td>
<td>81.4</td>
<td>75.0-100.0</td>
<td>.80</td>
</tr>
<tr>
<td>PM combined close and positive margin rate, % (n=675)</td>
<td>30.2</td>
<td>17.9-37.3</td>
<td>.02</td>
</tr>
<tr>
<td>Positive alone</td>
<td>11.0</td>
<td>7.1-14.3</td>
<td>.74</td>
</tr>
<tr>
<td>Operations attempting initial breast conservation, No. (n=675)</td>
<td>1.2</td>
<td>1.11-1.25</td>
<td>.50</td>
</tr>
</tbody>
</table>

Abbreviations: PM, partial mastectomy; TM, total mastectomy.

aPatients receiving primary systemic therapy prior to surgery were excluded.

bPatients with a history of ipsilateral breast cancer, irradiation, and multiple tumors in different quadrants were excluded.

Outcomes for axillary surgery are shown in Table 3. Both the number of nodes obtained during SLNB (P=.006) and the use of intraoperative node assessment (38.3%-100.0%; P<.001) varied among surgeons. The proportion of patients undergoing completion AD following a positive SLNB had a broad range (47.4%-100.0%) but did not achieve statistical significance (P=.20). The mean number of nodes following AD was 16.9 and did not vary among surgeons.

©2009 American Medical Association. All rights reserved.
Comparison of outcomes among the 3 high-volume surgeons demonstrated differences in the adjusted final TM rate (20.6%-34.2%; \( P = .03 \)), combined positive and close margin rate (26.4%-36.5%; \( P = .04 \)), and use of intraoperative sentinel lymph node assessment (38.3%-83.0%; \( P < .001 \)). There were no other significant differences noted. There was no significant variation in outcomes between surgeons within the low-volume group, but sample size (117 cases) may have limited the ability to detect differences.

**COMMENT**

Emphasis on quality assessment and quality improvement in health care has gained increasing importance in the conduct, monitoring, and payment for health care in the United States. This interest is shared among health care consumers (patients), payers (insurance, business leaders, and government), patient advocacy groups (Leapfrog [http://www.leapfroggroup.org]), and professional health care societies. Fueling this interest are innumerable reports that health care outcomes vary significantly among geographic regions, hospitals, and individual physicians. Health care costs in the United States greatly exceed those of other countries; furthermore, there is a growing demand for accountability regarding the value of health care received.8

Despite the frequency of breast cancer surgery worldwide, there remains a paucity of data for surgical outcomes discernable at the surgeon or hospital level. Birkmeyer et al9 have suggested that for commonly performed low-risk surgical procedures, outcome measures other than morbidity and mortality are necessary. A recent large-scale multicenter outcomes study evaluated morbidity and mortality following both PM and TM. Despite monitoring approximately 40 clinical variables, it was not possible to identify predictors of mortality, and only limited predictors of morbidity were identifiable.10 An accompanying editorial described these findings as underwhelming and "somewhat predictable."11 Adopting a similar quality-monitoring program on a broader basis would be costly, impractical, and likely untenable.

Outcome measures in our study were selected to reflect an important surgical decision (ie, initial PM or TM, reexcision following PM) or clinical decision (adjuvant therapy based on nodal metastases). As such, we would describe these as patient-centered outcomes. We also believed that these outcomes could be discretely identified and monitored at low cost and that they might vary among surgeons or hospitals treating a similar patient population. We excluded outcomes we felt were unlikely to affect a patient’s operative choice or selection of surgeon or hospital such as minute differences in wound infection rates, even if differences were real.10

Clear criteria for quality breast cancer operations are lacking, and debate on the selection of appropriate measures is ongoing.11-13 Our study offers an initial evaluation of the utility of several proposed quality measures. The clinical or cancer outcome of significance for which
each measure is a potential surrogate is listed in Table 5. We believe that some measures have enough established data regarding association with cancer outcomes, quality of life, or health care cost outcomes that they are suitable for adoption as quality measures. For others, the variation may reflect current uncertainty or evolving changes in breast cancer management, and clear links to outcomes of clinical importance have yet to be established. Nevertheless, variation in clinical care will likely be of keen interest to patients and may be important in evaluating health care costs.

In selecting the TM rate as a quality measure, we considered that PM (breast conservation) was feasible for most patients, was the less invasive and less life-altering procedure, and had long-term survival equivalent to that of TM.14,34,35 We recognized that TM is the more appropriate surgical therapy for certain patients.36 Previous studies using administrative databases have demonstrated wide variations in mastectomy rates across geographic ranges and patients of various socioeconomic backgrounds.37-39 These differences should be minimized in patients seen at a single center. Our institutional initial TM rate of 25.8% and adjusted final TM rate of 24.4% are well within the range of current series. Recent series have reported a TM rate of 35.8% in the United Kingdom37 and 37% at a US comprehensive cancer center.50 Our rate may be lower as all of the patients had a malignant diagnosis preoperatively and had initial surgery at our hospital. Our study did not demonstrate variation among surgeons for all of the patients but did show variation in the TM rate for T2 tumors (2.1-5.0 cm) and in the adjusted TM rate among high-volume surgeons. These differences are less likely to be explained by patient characteristics alone and may reflect differences in surgical management or opinions.

Mastectomy rates as a quality measure have limitations.41 A recent study demonstrated that well-informed patients using a decision aid chose mastectomy 35% of the time when presented with the options of TM or PM.16 Nevertheless, dramatic variation in mastectomy rates among hospitals or surgeons for patients with small tumors can reflect variation in processes of care, patient education and understanding, and surgeon’s influence.15 We propose that there should be an acceptable range of TM rates among surgeons treating a similar patient population. Recent European guidelines have suggested that breast conservation surgery should be achievable in 70% to 80% of all cases.3 That range appears achievable under appropriate circumstances, especially when diagnosis is confirmed before initial surgery and when the clinical tumor size is 2 cm or less. Appropriate granularity of data, ie, reporting TM rates by tumor size, will be necessary to compare outcomes and identify strategies to enhance care across regions of health care.

In selecting positive and close margin status following PM as a potential quality measure, we recognized that the

<table>
<thead>
<tr>
<th>Quality Measure</th>
<th>Institutional Average</th>
<th>Surgeon Average, Range</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lymph nodes obtained, No.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLNB (n=547)</td>
<td>3.7</td>
<td>2.7-4.2</td>
<td>.006</td>
</tr>
<tr>
<td>AD (n=200)</td>
<td>16.9</td>
<td>11.7-22.0</td>
<td>.15</td>
</tr>
<tr>
<td>Intraoperative lymph node assessment, % (n=566)</td>
<td>61.0</td>
<td>38.3-100.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>AD after positive SLNB, % (n=109)</td>
<td>59.6</td>
<td>47.4-100.0</td>
<td>.20</td>
</tr>
</tbody>
</table>

Abbreviations: AD, axillary dissection; SLNB, sentinel lymph node biopsy.

| Patients with a history of ipsilateral breast cancer were excluded. |

<table>
<thead>
<tr>
<th>Quality Measure</th>
<th>High</th>
<th>Low</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operations/y, mean, No.</td>
<td>57.4</td>
<td>7.9</td>
<td>NA</td>
</tr>
<tr>
<td>Time to surgery, mean, d4</td>
<td>22.0</td>
<td>25.6</td>
<td>.07</td>
</tr>
<tr>
<td>Adjusted initial TM rate, % (n = 827)b</td>
<td>21.3</td>
<td>24.5</td>
<td>.30</td>
</tr>
<tr>
<td>Adjusted final TM rate, % (n = 827)b</td>
<td>27.6</td>
<td>33.3</td>
<td>.20</td>
</tr>
<tr>
<td>PM combined close and positive margin rate, % (n = 675)</td>
<td>31.0</td>
<td>23.1</td>
<td>.15</td>
</tr>
<tr>
<td>Positive alone</td>
<td>11.3</td>
<td>9.0</td>
<td>.60</td>
</tr>
<tr>
<td>Operations following initial PM, mean, No. (n = 675)</td>
<td>1.2</td>
<td>1.1</td>
<td>.18</td>
</tr>
<tr>
<td>Lymph nodes obtained, mean, No. c</td>
<td>3.8</td>
<td>2.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SLNB (n=547)</td>
<td>17.3</td>
<td>14.8</td>
<td>.06</td>
</tr>
<tr>
<td>AD (n=200)</td>
<td>19.3</td>
<td>16.8</td>
<td>.06</td>
</tr>
<tr>
<td>Intraoperative lymph node assessment, % (n=566)</td>
<td>59.6</td>
<td>72.1</td>
<td>.06</td>
</tr>
<tr>
<td>AD after positive SLNB, % (n=109)</td>
<td>56.3</td>
<td>84.6</td>
<td>.04</td>
</tr>
</tbody>
</table>

Abbreviations: AD, axillary dissection; NA, not applicable; PM, partial mastectomy; SLNB, sentinel lymph node biopsy; TM, total mastectomy.

a Patients undergoing primary systemic therapy were excluded.
b Patients with a history of ipsilateral breast cancer, irradiation, and multiple tumors in different quadrants were excluded.
c Patients with a history of ipsilateral breast cancer were excluded.

Table 3. Outcomes of Axillary Surgery

Table 4. Breast Cancer Surgery Outcomes Comparing High- and Low-Volume Surgeons
The long-term clinical effect of variability in the volume of tissue excised. We did not record specimen volume, although the positive margin rate alone did not. Reasons for this variability may reflect technique differences or volume of tissue excised. We did not record specimen volumes, but this may be advantageous for a more robust comparison. The long-term clinical effect of variability in the combined positive and close margin rate in this series is unknown but is an area of future study. Feasibility of comparing institutions or surgeons regarding margins would require uniformity in margin assessment and reporting. This major principle in performing PM is complete excision of malignant tissue with an acceptable rim of surrounding normal breast tissue. This should be achieved in as few operations as possible while maintaining optimal cosmesis. Positive margins have clearly been shown to increase local recurrence rates, but there remains greater controversy around close margins (in both definition and management). We defined close margins as less than 1 mm. Many other institutions define close margins as less than 2 mm, and survey data from Europe suggest that some centers define close margins as less than 5 mm. We defined close margins as less than 1 mm. Many other institutions define close margins as less than 2 mm, and survey data from Europe suggest that some centers define close margins as less than 5 mm. The combined positive and close margin rate following initial PM in this series was 30.2% and varied among surgeons, although the positive margin rate alone did not. Reasons for this variability may reflect technique differences or volume of tissue excised. We did not record specimen volumes, but this may be advantageous for a more robust comparison. The long-term clinical effect of variability in the combined positive and close margin rate in this series is unknown but is an area of future study. Feasibility of comparing institutions or surgeons regarding margins would require uniformity in margin assessment and reporting. This outcome is controversial as a quality measure. Minimizing close and positive margins can be achieved by resecting a greater volume of tissue, thereby defeating the purpose of breast conservation. Nevertheless, the potential effects to patients both short-term (reexcisions, subsequent mastectomy, boost irradiation) and long-term (local recurrence) would suggest that margin status following initial PM is a reasonable quality measure. The reexcision rate following initial PM at our center was 17.5% and ranged from 10.7% to 20.1% among surgeons. Our rate is low compared with those of recent series from university hospitals, which have reported reexcision rates closer to 50%. We defined close margins as less than 1 mm. Many other institutions define close margins as less than 2 mm, and survey data from Europe suggest that some centers define close margins as less than 5 mm. The combined positive and close margin rate following initial PM in this series was 30.2% and varied among surgeons, although the positive margin rate alone did not. Reasons for this variability may reflect technique differences or volume of tissue excised. We did not record specimen volumes, but this may be advantageous for a more robust comparison. The long-term clinical effect of variability in the combined positive and close margin rate in this series is unknown but is an area of future study. Feasibility of comparing institutions or surgeons regarding margins would require uniformity in margin assessment and reporting. This outcome is controversial as a quality measure. Minimizing close and positive margins can be achieved by resecting a greater volume of tissue, thereby defeating the purpose of breast conservation. Nevertheless, the potential effects to patients both short-term (reexcisions, subsequent mastectomy, boost irradiation) and long-term (local recurrence) would suggest that margin status following initial PM is a reasonable quality measure. The reexcision rate following initial PM at our center was 17.5% and ranged from 10.7% to 20.1% among surgeons. Our rate is low compared with those of recent series from university hospitals, which have reported reexcision rates closer to 50%. This is consistent with the traditional National Surgical Adjuvant Breast and Bowel Project definition of negative margins. Variation in reexcision rates may reflect the surgeon’s positive and close margin rate as well as variable interpretation of the need to reexcise. Reexcisions are considerable burdens to patients in terms of time, inconvenience, discom-
fort, cosmesis, and psychological stress and may often lead to mastectomy. It is not unrealistic that 2- or 3-fold variation in reexcision rates may exist across hospitals, and the cost implications of such variation nationwide would be enormous. Such variation is worth identifying and opening a dialogue regarding operative techniques, patient selection factors in initial surgery, and perhaps more uniform selection criteria for use of neoadjuvant therapies. Recent European guidelines suggest that reoperation following initial surgery should not exceed 10%, but 20% seems to be a more reasonable benchmark.

Accurate axillary staging is integral to the selection of adjuvant therapy and identifying patients at higher risk for additional axillary disease. We hypothesized that variation might exist among surgeons in the management of the axilla both in treatment decision making and performance metrics. We demonstrated variation in the use of both intraoperative lymph node assessment and completion AD following positive SLNB at our center. This variation likely reflects current controversy regarding the necessity of AD, even in patients with positive SLNB considered at low risk for additional positive nodes being identified.

Our study is unique because outcome differences were not particularly striking between high- and low-volume surgeons. Factors contributing to this may be that 2 low-volume surgeons were fellowship trained in surgical oncology and were previously higher-volume breast surgeons. Significant variation in axillary measures was demonstrated among high-volume surgeons. This suggests that reporting of breast cancer surgical outcomes might be helpful on an individual surgeon basis rather than in broad categories (high volume vs low volume), which could mask important differences.

The quality measures proposed in this study vary significantly from the breast cancer care quality measures put forth by the National Cancer Quality Alliance in coordination with the American Society of Clinical Oncology and the American College of Surgeons (http://www.facs.org/cancer/qualitymeasures.html). Those measures focus on appropriate delivery of adjuvant irradiation, chemotherapy, and hormonal therapy. The measures put forth in this study are more likely to reflect differences in surgical judgment, bias, experience, technique, and performance. Variance in surgical practice patterns may prove to be as important in minimizing health care disparities as the more commonly highlighted differences in patient socioeconomic status or access to care. Even if the appropriateness of the measure as a quality indicator is not completely agreed on, monitoring surgical outcomes can offer additional benefits. Surgeons can provide patients with more accurate and realistic information based on their own individual performance. With appropriate confidentiality, quality monitoring will allow surgeons and institutions to evaluate their own individual performance in comparison with that of peers. This may afford opportunity to improve processes of care, enhance surgeon and patient education, and perhaps modify the conduct of actual breast cancer operations when desirable variation is identified. This hypothesis has been verified in the field of cardiac surgery, where significant improvements in outcomes were demonstrated based on responses to clinical outcomes data.

The physical, psychological, and economic implications of variation in the surgical practice of breast cancer surgery are staggering. Physicians and surgeons are at an important crossroads in defining appropriate measures of quality, other than mortality, for commonly performed surgical procedures. For cancer operations in particular, surgeons need to be involved in defining quality measures that matter to our patients. Looking beyond mortality (rare for most cancer operations) and compliance with Surgical Care Improvement Project measures (long-term cancer significance is likely unimportant) to measures that may affect local recurrence or even survival is needed. Careful and open comparisons of surgical outcomes can in the long run achieve improvements in the quality of cancer surgery and overall cancer care received by our patients. We feel that this study provides an initial step in defining quality measures for breast cancer surgery. Further investigation and consensus are required to determine appropriate national benchmarks for breast cancer surgical care.

Accepted for Publication: January 9, 2009.

Correspondence: Laurence E. McCahill, MD, Division of Surgical Oncology, Department of Surgery, University of Vermont, 89 Beaumont Ave, Given E309, Burlington, VT 05405 (laurence.mccahill@uvm.edu).

Author Contributions: Study concept and design: McCahill, Privette, and Ratliff. Acquisition of data: Privette, Sheehy-Jones, Ratliff, Majercik, Stanley, and Harlow. Analysis and interpretation of data: McCahill, Privette, James, Ratliff, Krag, and Harlow. Drafting of the manuscript: McCahill and Privette. Critical revision of the manuscript for important intellectual content: McCahill, James, Sheehy-Jones, Ratliff, Majercik, Krag, Stanley, and Harlow. Statistical analysis: Privette. Administrative, technical, and material support: McCahill, James, Sheehy-Jones, Ratliff, Krag, and Harlow. Study supervision: McCahill and James.

Financial Disclosure: None reported.

Previous Presentation: This paper was presented at the 89th Annual Meeting of the New England Surgical Society; September 27, 2008; Boston, Massachusetts; and is published after peer review and revision.

REFERENCES

McCahill et al provide us with an important first step in measuring multiple surgical parameters involving the treatment of primary invasive and noninvasive breast cancer. Their selection of the 8 quality outcome measures was based on literature as documented in their Table 5. This literature undoubtedly reflects a multidisciplinary diagnostic and therapeutic approach involving radiologists, pathologists, geneticists, medical and radiation oncologists, and both oncologic and reconstructive surgeons. It would be interesting to know how that approach was used in this work. In many high- or very low-volume practices it would be too onerous to have each new patient evaluated by the team, but in those cases standards of care are often adhered to by all involved in breast cancer care. Did such standards exist in the current study or was the team approach involved in breast cancer care? Did such standards of care are often adhered to by all low-volume practices it would be too onerous to have each new patient evaluated by the team, but in those cases standards of care are often adhered to by all involved in breast cancer care. Did such standards exist in the current study or was the team approach used with all of the patients?

The number of operations affects patients and the health care dollar. The decision for breast-conserving surgery certainly increases the risk for more operations. The role of magnetic resonance imaging in
