Variations in Persistent Use of Low-Value Breast Cancer Surgery

Ton Wang, MD, MS; Brooke C. Bredbeck, MD; Brandy Sinco, MS; Sarah Shubeck, MD, MS; Alison S. Baskin, BA; Ted Skolarus, MD, MPH; Lesly A. Dossett, MD, MPH

IMPORTANCE Through the Choosing Wisely campaign, surgical specialties identified 4 low-value breast cancer operations. Preliminary data suggest varying rates of deimplementation and have identified patient-level and clinician-level determinants of continued overuse. However, little information exists about facility-level variation or determinants of differential deimplementation.

OBJECTIVE To identify variation and determinants of persistent use of low-value breast cancer surgical care.

DESIGN, SETTING, AND PARTICIPANTS Retrospective cohort study in which reliability-adjusted facility rates of each procedure were calculated using random-intercept hierarchical logistic regression before and after evidence demonstrated that each procedure was unnecessary. The National Cancer Database is a prospective cancer registry of patients encompassing approximately 70% of all new cancer diagnoses from more than 1500 facilities in the United States. Data were analyzed from November 2019 to August 2020. The registry included women 18 years and older diagnosed as having breast cancer between 2004 and 2016 and meeting inclusion criteria for each Choosing Wisely recommendation.

MAIN OUTCOMES AND MEASURES Rate of each low-value breast cancer procedure based on facility type and breast cancer volume categories before and after the release of data supporting each procedure’s omission.

RESULTS The total cohort included 920,256 women with a median age of 63 years. Overall, 86% self-identified as White, 10% as Black, 3% as Asian, and 4.5% as Hispanic. Most women in this cohort were insured (51% private and 47% public), were living in a metropolitan or urban area (88% and 11%, respectively), and originated from the top half of income-earning households (65.5%). While there was significant deimplementation of axillary lymph node dissection and lumpectomy reoperation in response to guidelines supporting omission of these procedures, rates of contralateral prophylactic mastectomy and sentinel lymph node biopsy in older women increased during the study period. Academic research programs and high-volume facilities overall demonstrated the greatest reduction in use of these low-value procedures. There was significant interfacility variation for each low-value procedure. Facility-level axillary lymph node dissection rates ranged from 7% to 47%, lumpectomy reoperation rates ranged from 3% to 62%, contralateral prophylactic mastectomy rates ranged from 9% to 67%, and sentinel lymph node biopsy rates ranged from 25% to 97%. Pearson correlation coefficient for each combination of 2 of the 4 procedures was less than 0.11, suggesting that hospitals were not consistent in their deimplementation performance across all 4 procedures. Many were high outliers in one procedure but low outliers in another.

CONCLUSIONS AND RELEVANCE Interfacility variation demonstrates a performance gap and an opportunity for formal deimplementation efforts targeting each procedure. Several facility-level characteristics were associated with differential deimplementation and performance.

Published online February 3, 2021.

© 2021 American Medical Association. All rights reserved.
The provision of services without a clinically meaningful benefit is a national epidemic, costing the United States more than $100 billion dollars annually.1,2 Deimplementation is the science of eliminating low-value practices through evidenced-based processes.3–5 One prominent initiative to promote the deimplementation of low-value services is the Choosing Wisely campaign, which is a campaign by the American Board of Internal Medicine Foundation to identify unnecessary medical and surgical services.6 Seventeen surgical societies have participated in Choosing Wisely and identified 24 surgical procedures for deimplementation.7 However, despite general enthusiasm for reducing low-value surgery, clear gaps in evidence-based practice remain.

Examining the natural trend in de-escalation of surgical treatment for early-stage breast cancer offers a unique opportunity to identify determinants of deimplementation specific to low-value surgical procedures. Early-stage breast cancer is highly prevalent, carries an excellent prognosis, and multiple clinical trials support de-escalation of various treatments.8–10 Through Choosing Wisely, the American College of Surgeons, the Society for Surgical Oncology (SSO), and the American Society for Breast Surgeons have identified 4 low-value breast cancer treatments for elimination: (1) axillary lymph node dissection (ALND) for limited nodal disease in patients receiving lumpectomy and radiotherapy, (2) lumpectomy re-excision for close but negative margins for invasive cancer, (3) contralateral prophylactic mastectomy (CPM) in average-risk women with unilateral cancer, and (4) sentinel lymph node biopsy (SLNB) in clinically node-negative women 70 years and older with hormone receptor–positive (HR+) cancer.

Despite similar high-quality evidence supporting these recommendations, deimplementation has been inconsistent. Preliminary studies have shown a decrease in rates of ALND and lumpectomy re-excision at both institutional and national levels.2,11,12 In contrast, other studies suggest more than 80% of women 70 years and older with HR+ breast cancer receive SLNB2,13–15 and that CPM rates for patients with unilateral cancer are increasing.2,16 While previous studies have examined some tumor-level, patient-level, and clinician-level determinants of persistent use,5,17 to our knowledge, variation of deimplementation across facilities and procedures has not been described. Furthermore, the contribution of facility-level factors to variable use of unnecessary procedures has not been determined. Therefore, our aims are to (1) compare deimplementation rates and facility-level variation across procedures and (2) assess for facility-level determinants of deimplementation across procedures.

Methods

Data Source and Study Population

The National Cancer Database (NCDB) is based on hospital registry data collected from more than 1500 Commission on Cancer (CoC)-accredited facilities18 and captures approximately 70% of newly diagnosed cancer cases. All data are deidentified and Health Insurance Portability and Accountability Act-compliant. Because of the use of deidentified data, this study was deemed exempt by the University of Michigan institutional review board and patient consent was not obtained.

Using the NCDB, we identified women 18 years and older diagnosed as having breast cancer from 2004 to 2016. Four cohorts were created to evaluate ALND, lumpectomy margin re-reoperation, CPM, and SLNB rates. Women who received treatment outside of a CoC reporting facility or who received neoadjuvant systemic therapy were excluded. Detailed information regarding inclusion and exclusion criteria for each of the 4 low-value surgical targets is available in eAppendix 1 in the Supplement. Briefly, the ALND study cohort was based on the inclusion criteria of the American College of Surgeons Oncology Group Z0011 trial (n = 47174).9 The lumpectomy margin re-reoperation cohort was based on the 2014 SSO/American Society for Radiation Oncology (ASTRO) consensus statement of negative margin as “no tumor on ink” (n = 487443)19; re-reoperation was used as a proxy for re-excision similar to other NCDB studies.12,20 The CPM cohort included women with unilateral in situ or invasive stage 0 to II breast cancer who underwent mastectomy (n = 372561). The SLNB cohort included women 70 years and older with clinically node-negative, stage I to II, HR+ invasive breast cancer (n = 212733).21

Analysis of Procedure Variation at the Facility Level

We performed reliability adjustment using empirical Bayes methods to calculate hospital-level rates of each procedure from 2004 to 2016. To obtain an accurate estimate of facility performance, only hospitals with at least 10 patients per year were included. We created an interrupted time series hierarchical logistic regression model with an interrupted intercept, slope, and quadratic term, as well as a random intercept for facility, to account for clustering within facilities. We compared facility performance before and after publication of the data supporting omission of each procedure (2011 for ALND based on American College of Surgeons Oncology Group Z0011,9 2014 for lumpectomy re-reoperation based on the SSO/ASTRO consensus statement,19 2007 for CPM based on the SSO consensus statement that CPM for average risk patients with unilateral breast cancer is unnecessary,22 and 2013 for SLNB in women 70 years and older with HR+ breast cancer based on the 10-year CALGB 9343 results demonstrating that SLNB did not improve survival).23

Key Points

**Question** How do facility characteristics affect deimplementation of 4 low-value breast cancer operations in the Choosing Wisely campaign?

**Findings** In this cohort study in response to national recommendations to avoid 4 low-value procedures, use of 2 procedures decreased significantly and 2 other procedures increased in use. Academic research programs and high-volume facilities demonstrated the greatest reduction in use, with significant interfacility variation for each low-value procedure.

**Meaning** Facility-level characteristics were associated with use of low-value breast cancer operations.
We performed further analyses based on hospital volume and facility type. Hospital volume was based on average annual breast cancer case count and categorized as low (10-99 breast cancer cases), medium (100-199 cases), or high (≥200 cases). These cutoffs were chosen based on previously published ranges and to ensure an adequate number of facility-level and patient-level data across groups for multilevel analysis.24,25 A histogram detailing the proportion of hospitals and patients analyzed by hospital volume category for each procedure is provided in eAppendix 2 in the Supplement.

To determine hospital-level factors associated with successful deimplementation, we limited the sample to patients diagnosed from 2014 to 2016 when the maximum level of deimplementation could be expected. For each of the 4 procedures, reliability-adjusted hospital quintiles based on procedure rates were compared by cross-tabulation between surgical procedures using the Cochran-Armitage, Pearson χ², and Fisher exact tests as appropriate. We created a Pearson correlation matrix of reliability-adjusted rates to analyze whether trends existed between any 2 pairs of procedures.

A P value less than .05 was considered significant, and all P values were 2-sided. The P values represent differences in annual procedure rates or change in procedure rate over time using a hierarchal logistic model. All analyses were performed using SAS statistical software, version 9.4 (SAS Institute Inc).

Results

Trends in deimplementation of low-value breast cancer operations over time are shown in Figure 1. Consistent with prior data, we found ALND and lumpectomy reoperation
rates decreased rapidly in response to evidence demonstrating the safety of their omission. The ALND rates decreased from 63% (95% CI, 61.6-64.0) in 2004 to 14% (95% CI, 13.5-14.5) in 2016 (relative reduction of 78%). The greatest rate of change occurred from 2010 to 2011 (from 62% to 31%), corresponding to dissemination of Z0011 trial results (\( P < .001 \)).

Reoperation rates after lumpectomy decreased from 19% (95% CI, 18.8-19.7) in 2004 to 15% (95% CI, 14.4-15.1) in 2016 (relative reduction of 24%). The greatest rate of change was from 2013 (18%) to 2014 (16%), corresponding to the year of release of the SSO/ASTRO consensus statement designating a negative margin as “no tumor on ink” (\( P < .001 \)). For every 100 000 women, a 2% reduction corresponds to sparing 2000 women from the procedure.

By comparison, rates of CPM for patients with unilateral breast cancer and SLNB for older women have steadily increased since 2004. In 2016, 26% of women (95% CI, 24.7-25.9) with unilateral breast cancer undergoing mastectomy received CPM despite SSO guidelines in 2007 to avoid CPM for average-risk women,\(^2\) representing a nearly 2.5-fold increase since 2004 when the rate was 11% (95% CI, 10.3-11.2). The increase in CPM rates was statistically significant both before, in the year of, and after publication of the guidelines, with the greatest rate of change from 2007 to 2016 (from 16% to 26%; \( P < .001 \)). Similarly, rates of SLNB in women 70 years and older with clinically node-negative HR+ breast cancer increased from 78% (95% CI, 76.9-79.3) in 2004 to 87% (95% CI, 86.8-88.0) in 2012. The SLNB rates remained relatively stable from 2013 (88%) to 2016 (87%), despite evidence from the CALGB 9343 trial in 2013 showing no survival benefit.

Hospital-Level Variation in Rates of Low-Value Breast Cancer Operations
The hospital-level variation in rates of low-value breast cancer operations can be found in Figure 2. From 2014 to 2016, 389 hospitals performed at least 10 of each procedure per year with significant interfacility variation for each low-value procedure. Hospital-level ALND rates ranged from 7% to 47% (mean [SD], 17.8 [5.8]), lumpectomy reoperation rates ranged from 3% to 62% (mean [SD], 16.8 [7.0]), CPM rates ranged from 9% to 67% (mean [SD], 31.8 [11.3]), and SLNB rates ranged from 25% to 97% (mean [SD], 85.2 [9.0]). The ranked order of adjusted rates across hospitals is displayed in caterpillar plots. The maximum Pearson correlation coefficient between any 2 of 4 procedures was less than 0.11, suggesting little correlation in hospital performance across procedures. For ALND and lumpectomy reoperation, hospitals in the top vs bottom quintile did not differ based on breast cancer volume or facility type. However, significantly more integrated network cancer programs were in the highest quintile for CPM rates compared with community cancer programs (23% vs 2%; \( P < .001 \)). Additionally, more comprehensive community cancer programs were in the highest quintile for SLNB rates for women 70 years and older with HR+ breast cancer compared with academic research programs (48% vs 21%; \( P = .05 \)). Further detail is provided in eAppendix 3 in the Supplement.

Trends in Performance of Low-Value Breast Cancer Operations Based on Facility Type
Trends in the performance of low-value breast cancer operations based on facility type can be found in Figure 3. For rates of ALND, deimplementation ranged from a relative reduction of 57% (42% in 2004 to 18% in 2016) for community cancer programs to a relative reduction of 79% (68% in 2004 to 15% in 2016) for academic research programs (\( P = .055 \) for academic vs nonacademic). Academic research programs were the only facility type to have a significant decrease in ALND rates prior to Z0011 (from 68% to 65%; \( P < .001 \)). From 2010 to 2011 and from 2011 to 2016, all facility types except for community cancer programs had a significant decrease in ALND rates. For lumpectomy reoperation, deimplementation ranged from a relative reduction of 20% (19% in 2004 to 15% in 2016) for community cancer programs to a relative reduction of 30% (22% in 2004 to 15% in 2016) for academic research programs (\( P = .002 \) for academic vs nonacademic). The trend in deimplementation occurred prior to the release of the SSO/ASTRO guidelines in 2014 with a significant decrease in lumpectomy reoperation for every facility type except for community cancer programs. During the year of guideline release from 2013 to 2014, there was a significant decrease in rates for every facility type. Rates decreased from 19% to 16% in community cancer programs, from 19% to 17% in comprehensive community cancer programs, from 19% to 17% in academic research programs, and from 19% to 17% in integrated network cancer programs (\( P < .001 \)). For CPM, academic research programs had the lowest rates at the end of the study period (26% in 2016) while integrated network cancer programs had the highest (32% in 2016; \( P < .001 \) for academic vs nonacademic). Similarly, academic research programs had significantly lower rates of SLNB in older women with HR+ breast cancer at the end of the study period (84% in 2016) compared with all other facility types (87%-88% in 2016; \( P < .001 \) for academic vs nonacademic). Further detail is provided in eAppendix 4 in the Supplement.

Trends in the performance of low-value breast cancer operations based on facility annual breast cancer volume can be found in Figure 4. High-volume hospitals had the greatest decrease in ALND rates (from 65% in 2004 to 14% in 2016; relative reduction, 79%). Medium-volume hospitals had the smallest decrease in ALND rates and the highest ALND rates at the end of the study period (from 62% in 2004 to 17% in 2016; relative reduction 72%; \( P < .001 \) for high- vs medium-volume facilities). Whereas ALND rates only decreased at high-volume facilities before 2011, all facilities had significant decreases in ALND rates after Z0011. For lumpectomy reoperation, high-volume hospitals had the greatest reduction but also had the highest reoperation rates at the end of the study period (from 22% in 2004 to 16% in 2016; relative reduction, 28%; \( P = .002 \) for high- vs medium-volume facilities). All hospitals by volume category de-
creased reoperation rates during the year of guideline release. Rates decreased from 18% to 15% in low-volume facilities, from 18% to 16% in medium-volume facilities, and from 21% to 18% in high-volume facilities ($P < .001$). Although CPM rates increased for all facilities regardless of hospital volume, low-volume hospitals had the lowest rates of CPM during the study period (from 7% in 2004 to 24% in 2016), whereas high-volume facilities had the highest rates of CPM (from 14% in 2004 to 29% in 2016; $P < .001$ for high- vs low-volume facilities). Conversely, high-volume hospitals had the lowest rates of SLNB during the study period (from 77% in 2004 to 86% in 2016), while low-volume hospitals had the highest rates (from 79% in 2004 to 88% in 2016; $P = .04$ for high- vs low-volume facility rates in 2016). Further detail is provided in eAppendix 5 in the Supplement.

Discussion

To our knowledge, this is the first study to explore facility-level variation and determinants of differential deimplementation of low-value surgery in a single disease. We identify 3 findings to inform future efforts to reduce overtreatment. First, natural deimplementation of the 4 procedures was variable despite similar levels of evidence supporting treatment de-escalation. Second, we demonstrated significant interfacility variation in deimplementation; being a positive outlier of deimplementation for some procedures did not translate to being a positive outlier for others. Finally, several facility-level characteristics were associated with deimplementation performance, suggesting strategies...
to reduce low-value care can be tailored to institutional factors.

This study confirms prior data demonstrating that deimplementation of ALND and lumpectomy reoperation occurred rapidly after published evidence supported omission of these procedures. This finding is notable considering it takes 17 years, at least historically, for clinical practice to change in light of research findings. Conversely, SLNB rates in older women with early-stage HR+ breast cancer and CPM rates in average-risk women with unilateral breast cancer have steadily increased since 2004. Thus, understanding the factors contributing to the early deimplementation of ALND and lumpectomy reoperation may help identify strategies to reduce other low-value surgical procedures performed at persistently high rates.

In prior qualitative work, we and others identified patient and clinician-related factors facilitating deimplementation. For ALND, both surgeons and patients viewed lymphedema risk as significant, which likely contributes to eagerness to omit ALND. Notably, this finding is mirrored by the rapid deimplementation of completion lymph node dissection in patients with melanoma after the Multicenter Selective Lymphadenectomy Trial-II (MSLT-II) demonstrated no overall survival benefit. Additionally, clinicians have cited the strong evidence base and wide dissemination of the Z0011 trial for ALND and SSO/ASTRO recommendation of a negative margin as “no tumor on ink” as reasons why these practices have decreased. This is supported by our study, which found considerable variation in ALND and lumpectomy reoperation rates by facility type before the respective guidelines but minimal variation afterwards, suggesting these recommendations provided clarity on appropriate indications for ALND and lumpectomy reexcision where there had previously been uncertainty. In contrast, qualitative studies found surgeons are not convinced about the quality of the evidence supporting SLNB omission, are not familiar with national recommendations to avoid SLNB, and feel the procedure adds minimal time and risk to a patient’s operation. This suggests quantifying and communicating overtreatment harms to clinicians and patients (eg, the care cascades associated with unnecessary SLNB or financial...
toxicity and increased risk of complications associated with CPM or unindicated lumpectomy re-excision) may be effective strategies for deimplementation.

While factors contributing to the dramatic increase in CPM rates over the last decade have been investigated, nearly all efforts to eliminate this low-value practice have focused on demand-side, or patient-level, factors. Contralateral prophylactic mastectomy has traditionally been viewed as a patient preference-sensitive procedure, and high-procedure rates are attributed to patient-level factors including younger age, insurance status, desire for peace of mind, fear of recurrent disease, and misperceptions about its influence on survival rates. However, our study shows a direct association with hospital volume and facility type, suggesting significant supply-side contributions to this trend. The finding that high-volume hospitals have the highest CPM rates likely reflects wider availability of breast reconstruction, which is known to be strongly correlated. However, despite having access to breast reconstruction and likely high-volume breast cancer surgery practices, academic research facilities had the lowest rates of CPM at the conclusion of the study period.

Importantly, our study demonstrates facilities are inconsistent in deimplementation performance, suggesting that reducing overtreatment is not an inherent trait associated with a particular facility. Given these findings, strategies for deimplementation should target each procedure individually, with attention to the varied stakeholders involved. One potential approach is to develop a deimplementation toolkit adapted to an individual hospital’s performance across multiple metrics because not all facilities will require the same interventions. Although hospitals are not consistently positive or negative outliers in deimplementation performance, there are some key differences in deimplementation based on facility type and volume. Facility characteristics play a key role in eliminating overtreatment through organizational culture, leadership, and resources. Recognizing and targeting specific facility-level
facets is an attractive strategy for reducing low-value services because implementing change in health care frequently occurs at the hospital-level. Academic research programs had the lowest rates of CPM and SLNB in older women at the end of the study period, and by some measures were the most successful in deimplementing ALND. Additionally, high-volume facilities had the lowest rate of each low-value procedure except for CPM.

Some of the trends noted in our study may reflect the diversity of clinicians caring for patients with breast cancer. Unlike some cancers that are largely centralized to academic medical centers, most patients with breast cancer are treated in nonacademic settings, where there is variation in clinician training and procedural volume. Whereas specially-trained or high-volume surgeons may be more comfortable omitting therapies, clinicians with less oncologic experience may be more aggressive owing to concerns about errors from omission. Studies supporting this hypothesis have associated higher-volume breast surgeons with implementation of national quality metrics including high rates of breast-conserving surgery, oncoplastic surgery, and improved patient satisfaction. Additionally, high-volume breast surgeons may concentrate in academic facilities, whose resources, culture, and payment structure encourage multidisciplinary and evidence-based care.

Limitations

While major strengths of the NCDB include breadth of the patient population and facility-level data, it is limited by its retrospective nature and available variables. As a result, eligible patient cohorts were based on surrogate measures available through NCDB (eg, the use of reoperation rather than re-excision without definitive knowledge of the lumpectomy margin status). However, methods in this study have been used in prior literature. There may be important facility-level characteristics contributing to use of low-value services not represented in this data set (eg, reimbursement structures). Because the NCDB is composed of CoC facilities, this database may be skewed toward more complex diagnoses. However, the NCDB is a comprehensive database that encompasses approximately 70% of patients diagnosed with breast cancer, and we would expect that CoC hospitals are excellent targets for assessing gaps in evidence-based practice. Increased genetic testing (which is not an available variable through NCDB) may result in higher CPM rates, but the prevalence of pathogenic germline mutations is low and is unlikely to account for these increases alone. Finally, we note that an ideal rate of deimplementation has not been established and may vary by procedure. Even in the case of ALND, where surgeons and patients recognize the significant complication risks, 15% to 20% of patients continue to undergo ALND. This is likely owing to factors that cannot be quantified by a large data set such as the NCDB.

Conclusions

Despite similar evidence and national recommendations supporting the omission of 4 low-value breast cancer procedures, only 2 have been successfully deimplemented. Several facility-level characteristics were associated with deimplementation performance, with academic research facilities and facilities with a high volume of patients with breast cancer demonstrating the greatest reduction in use of these low-value procedures. However, hospitals were not uniform in their deimplementation performance across all 4 procedures, suggesting that success at reducing overtreatment is not an inherent trait associated with a particular hospital. Significant interfacility variation demonstrates a performance gap for many centers and room for formal deimplementation efforts targeting each procedure.
Variations in Persistent Use of Low-Value Breast Cancer Surgery

Original Investigation Research


Boughey JC, Haffty BG, Habermann EB, Hoskin TL, Goetz MP. Has the time come to stop surgical staging of the axilla for all women age 70 years or older with hormone receptor-positive breast cancer? Ann Surg Oncol. 2017;24(3):654-677. doi:10.1245/s10434-016-5740-z


was ineffective. Health Aff (Millwood). 2016;35(7):1309-1315.


