Protocol

For reprint orders, please contact: info@becaris.com

Journal of **Comparative Effectiveness Research**

A protocol for three observational cohort studies evaluating adverse outcomes, excess costs and repeat procedures after surgery for breast cancer in the USA

Sam E Wing*[,1], Yuki Liu[1], Jonathan C Johnson[2], Kristin J Moore[2] & Feibi Zheng[1,3]

[1]Intuitive Surgical, Sunnyvale, CA 94086, USA
[2]Optum Inc, Eden Prairie, MN 55344, USA
[3]Michael DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA
*Author for correspondence: Tel.: +1 206 669 7211; sam.wing@intusurg.com

Aim: Surgical therapy for early-stage breast cancer, including lumpectomy and mastectomy, are common treatments for early-stage breast cancer. Despite having favorable survival outcomes, these procedures can lead to repeat surgeries, adverse outcomes, excess costs and potentially aggressive resections. This is a protocol for a study aims to evaluate three main areas: the risk factors, costs and complications of re-operations following lumpectomy ('Lumpectomy Cohort'), the identification of patients potentially overtreated with mastectomy ('Mastectomy Cohort') and the cost and healthcare resource utilization associated with nipple necrosis following nipple-sparing mastectomy (NSM) ('Nipple Necrosis Cohort'). **Materials & methods:** A retrospective cohort analysis will be conducted using Optum's de-identified Market Clarity Data (2007–2023), which integrates medical and pharmacy claims with electronic health records. Patients will be included based on specific procedure and diagnosis codes, with additional data extracted from unstructured clinical notes using natural language processing. The study will analyze patient demographics, baseline health, surgical details, and outcomes, including costs, complications, reoperations and mortality. Data will be analyzed descriptively, with Kaplan–Meier analyses for time-to-event outcomes and Wilcoxon Signed Rank tests for cost comparisons. **Results:** Preliminary cohorts are expected to include 26,472 lumpectomy patients, 16,836 mastectomy patients and 6828 NSM patients with 541 cases of nipple necrosis. **Conclusion:** This study will provide comprehensive insights into the patient journey – highlighting the costs and patient outcomes following lumpectomy, mastectomy and NSM – potentially guiding better clinical decision-making and resource allocation.

Plain language summary

What is this article about? This article describes a plan for a study that looks at the outcomes and costs of different surgeries for early-stage breast cancer. It focuses on three main areas: the reasons and costs for needing more surgeries after a lumpectomy, identifying demographic and clinical profiles of patients undergoing mastectomies and the costs and problems related to nipple damage after nipple-sparing mastectomies.

What methodology will be followed? The study will use past medical records and data from 2007 to 2023, utilizing information from medical insurance claims and electronic health records. The study will look at patient details, health status, surgery information and outcomes like costs and complications.

Why is this important? We aim to provide a better understanding of the patient experience and identify areas where care can be improved for those undergoing breast cancer surgeries. By learning about the factors and costs associated with additional surgeries, aggressive resections and complications, the findings could help doctors make better decisions and use resources more effectively.

First draft submitted: 7 February 2025; Accepted for publication: 9 June 2025; Published online: 14 July 2025

Keywords: breast cancer • breast surgery • lumpectomy • mastectomy • nipple-sparing mastectomy • NSM • real-world data analysis • RWD analysis

In the USA, there are approximately 275,000 new breast cancer diagnoses and 43,000 breast cancer related deaths each year [1]. Diagnosis most frequently occurs at an early stage (e.g., stage I or II), comprising about 60% of patients [2]. Among these early-stage patients, 95% go on to receive some form of surgical therapeutic intervention, with 61% receiving breast conserving surgery (BCS, i.e., lumpectomy with radiation) and 34% receiving a mastectomy. Trends for utilization of lumpectomy relative to mastectomy are increasing in the USA since the early 2010s [3], though mastectomy remains a primary surgical option for early-stage patients who may have contraindications for lumpectomy and radiation. Contraindications may include BRCA1 and/or 2 gene mutation carriers, previous chest radiation recipients, disease multicentricity and connective tissue disorders [4]. Among mastectomy patients, almost 15% receive a nipple-sparing mastectomy (NSM) [5], which preserves the nipple-areolar complex (NAC) and can improve cosmetic satisfaction and quality of life compared with total mastectomies [6].

Survival of early-stage breast cancer is generally high, with 5-year survival rates of >99% for stage I and 93% for stage II patients [7]. However, the patient journey may be highly suboptimal for individuals both in the lumpectomy and mastectomy pathways. Unplanned re-operation after a lumpectomy is not uncommon, as approximately one in four of lumpectomy patients go on to have a second lumpectomy or subsequent mastectomy after their index surgery [8]. Common reasons for re-operations include the detection of positive surgical margins or disease recurrence [9,10]. These recurrences can negatively impact patient quality of life [11] and reoperation contribute high additional costs to the financial impact of treating breast cancer [12].

Mastectomy may be the preferred surgical intervention for early-stage patients with the previously mentioned oncologic contraindications. Logistical considerations may also influence this surgical selection, as women living further away from medical centers are more likely to receive a mastectomy compared with lumpectomy, which is frequently paired with postoperative radiation [13]. While oncologic outcomes are likely to be mostly comparable between patients who undergo mastectomy compared with lumpectomy with adjuvant radiotherapy [14], mastectomy patients have reported lower quality of life compared with lumpectomy patients [15]. Further, postoperative complication rates among mastectomy patients are around 10%, including both local (e.g., surgical site infection and flap failure) and systemic (e.g., pneumonia and urinary tract infections) complications [16]. Two-year post-lumpectomy complication rates (e.g., infection, hematoma and/or seroma, breast pain, etc.) occur in approximately 30% of patients which can lead to excess average costs to payors ranging from $549 to $10,402 per patient [17].

For patients undergoing an NSM, the surgical approach required for preservation of the NAC may also endanger the preservation of adequate blood supply to the same region and can result in necrosis. This adverse outcome has been previously reported in up to 29% of NSM cases, though most studies report a nipple necrosis rate closer to approximately 10% [18–25]. Cases of nipple necrosis can require significant interventions depending on complication severity, including complete removal of the NAC.

Aims & objectives

We have three main objectives for these real-world data (RWD) studies. First, we aim to study the patient population undergoing index lumpectomy procedures for the treatment of early-stage breast cancer in order to better understand the patients who are at highest risk for a re-operation and to quantify the cost and consequences of their subsequent care ('Lumpectomy Cohort'). Examining patient risk profiles among patients with re-operations will help describe how contraindication guidelines [26] have diffused throughout clinical practices in the country, as well as how risk mitigation techniques (i.e., localization and intraoperative frozen section review) have been adopted. Further, we will be able to describe the additional burden to the patient and healthcare system incurred by the requirement of additional surgical procedures. This will include the prevalence and cost of additional repeated breast surgeries, as well as any additional therapies (e.g., radiation, chemotherapy, or immunotherapy), imaging (e.g., CT or MRI scans), and other healthcare resource utilization. Second, we aim to study the population of patients undergoing mastectomies for early-stage cancer to identify the disease and patient characteristics of individuals that underwent this most aggressive surgical option ('Mastectomy Cohort'). While certain patients are clinically indicated for this more aggressive intervention, identification of patients who are 'overtreated' with mastectomy could help quantify the proportion of patients who could be well served by a less aggressive surgical option. Third, we aim to quantify the cost and consequences of nipple necrosis following an NSM ('Nipple Necrosis Cohort'). While some mild cases

may resolve with minimal interventions, more severe cases may require very resource-intensive care. This analysis will help define the range of treatment healthcare resource utilization associated with nipple necrosis cases and the additional financial burden it places on patients and payers.

Materials & methods

Data source

This study will be a retrospective cohort analysis of adult early breast cancer patients using administrative claims, enrollment information, and linked clinical data from Optum's de-identified Market Clarity Data (2007–2023). Optum® Market Clarity Data is an integrated, multi-source dataset comprising medical claims, pharmacy claims, and electronic health records. Optum Market Clarity links electronic health record data – including lab results, vital signs, measurements, diagnoses, procedures, and information derived from unstructured clinical notes using natural language processing – with historical, linked administrative claim data, including pharmacy claims, physician claims, clinical information facility claims, and medications prescribed and administered. Optum Market Clarity is statistically de-identified under the HIPAA Privacy Rule's Expert Determination method and managed according to Optum customer data use agreements (27,28). Multi-source medical claims data integrates administrative medical and pharmacy claims, along with linked enrollment information. This dataset encompasses a broad spectrum of US insurance types, including commercial plans, Medicare Advantage, Medicaid, and others. Administrative medical data is sourced from both inpatient and outpatient electronic health records (EHRs) across the US. These records are processed, normalized, and standardized to ensure consistency. The structured EHR data includes demographics, prescribed medications, immunizations, contraindications, lab results (including microbiology), vital signs, clinical observations, inpatient stay details, and coded diagnoses and procedures. In addition, Optum employs advanced natural language processing (NLP) technology to extract clinical concepts from unstructured physician notes. This enhances the dataset with valuable insights such as signs and symptoms, family medical history, biomarker data, and oncology-specific variables – many of which may not be explicitly recorded in structured EHR fields. In total, Market Clarity contains data provided from 150 different commercial and government insurance payors on over 70 million patients [29].

The EHR data are compiled into both structured and unstructured data fields. Structured data fields include common variables like patient characteristics (e.g., BMI, race/ethnicity and smoking history), clinical details (e.g., laboratory results and diagnoses) and prevalidated cancer-specific characteristics (e.g., cancer stage and BRCA1/2 mutation status). Some patient characteristics, like race, ethnicity and smoking status were self-reported. The unstructured data are comprised of physician free text documentation, which are generated during the course of patient care, including office visits and operative notes. Public involvement is not planned as part of this database analysis.

Study cohort & procedures

Patients will be included if they appear in the Optum Market Clarity dataset from 1/1/2012 through 12/31/2022. All relevant procedures and diagnoses will be identified using International Classification of Diseases 9th revision (ICD-9), ICD-10, Current Procedural Terminology and Healthcare Common Procedure Coding System codes (Supplementary Table 1) [30–32]. For procedures and diagnoses without a clear coding framework or guidance, we will query the unstructured provider notes with relevant keywords to identify patients of interest. Breast cancer patients will be identified based on the presence of ICD-9 or ICD-10 diagnosis codes and will be excluded in the presence of evidence of metastatic disease. Lumpectomy and mastectomy patients will be identified using ICD procedure and Current Procedural Terminology codes. Males or patients with a diagnosis of a male breast cancer are excluded from analysis as male breast cancer may have unique disease characteristics and, as such, have treatment pathways that differ from female patients [33].

To be further considered for inclusion, patients must have at least two nondiagnostic medical claims with a diagnosis code for breast cancer on two separate dates during the study period. The earliest breast cancer claim during this period will be considered as the diagnosis date. Once the breast cancer cohort has been identified, patients will be assigned to both the lumpectomy and mastectomy cohorts based on receipt of either lumpectomy or mastectomy. If a patient received both a lumpectomy and mastectomy and is eligible for inclusion, they will be assigned to both the lumpectomy and mastectomy study cohorts. All patients must have 6 months of continuous insurance enrollment prior to the index, with 12–60 months of postoperative continuous enrollment in the

10.57264/cer-2025-0013

lumpectomy cohort – allowing for sufficient follow-up to detect postlumpectomy re-operations – and 6 months in the mastectomy cohort – allowing for sufficient follow-up to detect postmastectomy complications.

Assignment to the nipple necrosis cohort will be defined by receipt of an NSM. NSM does not have a clearly defined procedure code in the US, so we will utilize the unstructured physician notes to define this cohort. As above, patients will be required to have the same diagnosis codes for inclusion, as well as a mastectomy procedure code. Patients then must have the presence of physician notes indicating receipt of a NSM (text string queries in Supplementary Table 2). Nipple necrosis cases will be further defined from the set of patients who were identified as NSM patients and then had the presence of a NAC ischemia/necrosis text string search in their free text notes (text string queries in Supplementary Table 2). Patients will be excluded from the mastectomy cohort if they are identified as NSM patients.

After patients are allocated into the different analytical cohorts, separate index dates and inclusion/exclusion criteria will be applied as previously described. The index date of each analysis will be the earliest claim for that specific breast cancer surgery procedure (lumpectomy or mastectomy). Patients will be followed until the earliest of death, end of continuous enrollment, or end of the study period. Data will be analyzed descriptively, accounting for the variable follow-up period as necessary.

Demographic information

Patient demographics will be identified from insurance enrollment information. These will include age, gender, race, ethnicity, insurance type (e.g., commercial, Medicare, etc.) and geographic region. Geographic region will be reported in alignment with one of five of the US Census Regions (i.e., Northeast, Midwest, South, West and other).

Baseline health, index surgical encounter, outcomes

Baseline health will be comprised of comorbidities, BMI, smoking status, medication use, laboratory results and clinical details about the breast cancer. Comorbidities evaluated will include both those comorbidities used to construct both the Elixhauser Comorbidity and Charlson Comorbidity Indexes [34,35]. As nipple necrosis could be related to poor vascularization, the nipple necrosis cohort will also identify conditions that could impact vasculature, including the presence of diabetes and connective tissue disorders. Patients with missing data on covariates will not be excluded and the prevalence of missing data will be reported.

BMI is derived from structured data fields in the EHR, which can either be stored as a precalculated BMI field or calculated from a patient's height and weight. BMI will also be categorized as underweight (0–18.5), low normal (18.5–25), overweight (25–30) and obese (greater than 30). Smoking status will also come from structured EHR data and will be defined as current, never or former smoker. For both smoking and BMI, their status will be selected from observations on or before the index surgical date, selecting the closest available observation to the index surgical date.

Use of immunologic, anticoagulant, steroids, antidiabetic and GLP-1 agonist medications during the baseline period will be indicated and identified from claims data using National Drug Codes. Laboratory results for HbA1c, platelet count, white blood cell count, albumin, prothrombin time test, cotinine and thyroid stimulating hormone will be extracted from EHR data. These results will be selected from tests administered on or before and closest to the index surgical date.

Disease details are collected from the EHR, of which some are stored in prebuilt structured fields and others are collected from unstructured notes and converted into structured fields using validated NLP algorithms. The NLP process begins with a random sample of data that is double annotated by two annotators with conflicts being resolved by a third review by a curator. This data becomes the gold standard which is partitioned into training, validation and testing subsets for NLP model development. Supervised machine learning models are trained to identify broader patterns not explicitly and manually created to enable the system to generalize to relevant contexts. The models are designed to identify the occurrences of the desired oncology concepts, their relations and the surrounding contexts such as temporal status and assertion (qualifiers) and apply labeling predictions to incoming data. The models are evaluated against the testing data and performance is quantified using precision, recall and F1 (the harmonic mean of precision and recall). Once models are finalized, they are run at scale, extracted concepts are normalized, de-duplicated and verified before being integrated into the database. Prebuilt disease details will include tumor grade, tumor size, presence of microcalcifications, nipple involvement and tumor location. NLP-derived disease details will include estrogen receptor/progesterone receptor (ER/PR) status and BRCA mutation status.

Information about the index surgical encounter will include procedure details, peri- and postoperative complications and surgeon/hospital profiles. Use of intra-operative adjuncts (i.e., needle localization, specimen radiograph and/or frozen pathology section review) and surgical margins will provide context about the index surgical encounter. Complications to be collected include hematoma, seroma, infections, hemorrhage, wound dehiscence, lymphedema, nerve injury, necrosis and implant loss. Surgeon specialty (e.g., breast surgeon, general surgeon, etc.) and the hospital's number of beds, and community versus academic center status will additionally be reported.

Patient outcomes will be derived from a combination of claims and EHR data. Main outcomes to be identified include mortality, postoperative procedure and costs. Data on mortality is captured from a variety of sources in order to provide the most complete information on death available including the Social Security Administration SSA public use Death Master File (SSADF). Postoperative procedures for the lumpectomy cohort include the performance of a repeat lumpectomy or subsequent mastectomy after the index lumpectomy. The mastectomy cohort will also identify subsequent cases breast reconstruction and implant removal. The nipple necrosis cohort will quantify the necrosis treatments (e.g., hyperbaric oxygen therapy, debridement and skin graft) after an NSM. All costs during the follow-up period will be standardized and stratified into pharmacy and medical costs. Medical costs will be further divided into costs incurred in the ambulatory, inpatient, emergency and other settings. More details on variables to be analyzed are available in Supplementary Table 3.

Statistical analysis plan

In all cohorts, patient demographics and clinical characteristics will be described for patients undergoing lumpectomies, mastectomies and NSM using means for continuous variables and percentages for categorical variables. These will further be compared between subgroups (e.g., patients with and without repeat procedures in the lumpectomy cohort or patients with and without nipple necrosis in the Nipple Necrosis cohort) using *t*-tests for continuous variables and chi-square tests for categorical variables. Among patients in the lumpectomy cohort, prevalence and time to repeated lumpectomy or subsequent mastectomy will be calculated, with 5-years of follow-up. 4- to 1-year follow-up times will also be presented in sensitivity analyses. Time-to-event (i.e., additional repeated surgery) also be examined using Kaplan–Meier analyses. Patient survival (i.e., no repeated surgery), including median and interquartile range, will be reported at 3, 6, 9, 12, 24, 36, 48 and 60 month time points. Log-rank tests will be used to identify differences in survival between demographic (e.g., age and race/ethnicity) and clinical subgroups (e.g., tumor grade and ER/PR status). Patients will be classified by the total number and combination of surgeries received (e.g., one lumpectomy only vs one lumpectomy and one repeat lumpectomy vs one lumpectomy and one subsequent mastectomy). Median healthcare costs, including the previously mentioned cost subgroups, will be compared between the groups, using Wilcoxon Signed Rank tests to identify differences. p-values less than 0.05 will be considered statistically significant and all tests will be two-sided.

Patients in the mastectomy cohort will be evaluated to construct an overall patient journey, including pre-, peri-, and post-operative descriptive analyses. In addition to baseline health variables, preoperative characteristics will include receipt of neoadjuvant therapy and previous breast surgery. Perioperative characteristics will include the distribution of surgeon specialties, hospital sizes and types and complications. Post-operative characteristics will include total healthcare cost, breast reconstruction patterns and 5-year all-cause mortality for the lumpectomy cohort and 1-year all-cause mortality for the mastectomy cohort.

Nipple necrosis patients will be analyzed to describe the cost and consequences of this adverse outcome. We will describe the frequency and distribution of interventions required to resolve the necrosis and summarize the median cost these interventions. We will also conduct a similar analysis of a larger subset of patients who have had either a nipple or skin necrosis post-operatively, to evaluate the impact of broader necrotic events on healthcare resource utilization.

Preliminary analytical cohorts

Based on the data source and inclusion/exclusion criteria outlined above, the analytical cohorts are expected to contain the following number of participants:

- Lumpectomy cohort: 26,472 patients with at least 5 years of continuous enrollment follow-up (101,908 patients with at least 1 year of continuous enrollment follow-up for sensitivity analyses).
- Mastectomy cohort: 16,836 patients with at least 5 years of continuous enrollment follow-up (62,707 patients with at least 1 year of continuous enrollment follow-up for sensitivity analyses).

10.57264/cer-2025-0013

- Nipple necrosis cohort: 6828 NSM patients, with 541 nipple necrosis patients (1233 patients with skin or nipple necrosis for sensitivity analyses).

Strengths & limitations

The Market Clarity database is a robust analytical resource, containing claims and medical record data from a wide range of payors and providers across the country. This allows for the analysis of RWD that reflects practice patterns of a diverse set of patients and physicians. Such practice patterns may be challenging to detect in smaller clinical databases that contain information from a single surgeon or hospital. Additionally, linking claims information with medical records allows for the impact of the suboptimized patient journeys to be considered from the perspective of multiple stakeholders, including added burden to the healthcare system and financial impact to payors.

The use of RWD and this database does have limitations. The claims and medical record data used for the curation of these analyses were not collected for the intention of doing research. This could potentially lead to inconsistent reporting or the absence of healthcare interactions that occurred outside the provider and payor systems from which these data are gathered. Further, these claims data are subject to coding errors and the common practice of requiring multiple diagnosis codes on separate dates and/or treatment, which could lead to the misclassification of both procedures and diagnoses, as well as their respective dates. This data also may not be representative of the broader population, especially those without insurance coverage or without access to healthcare. No data on patient preferences or patient reported outcomes were available in this data, which may mask part of the decision-making process and quality of life outcomes. Finally, these observational data also are not randomized and are subject to potential confounding and other biases. However, these weaknesses – often inherent to observational data – are offset by the practicality and generalization enabled by utilization of a broad, RWD source.

Summary points

- The patient journey for recipients of breast surgery for the treatment of cancer is often not optimized, leading to frequent and costly reoperations and complications.
- The study aims to analyze real-world data (RWD) to understand patient outcomes and healthcare costs associated with different surgical interventions for early-stage breast cancer, focusing on lumpectomy, mastectomy and nipple-sparing mastectomy.
- This is a retrospective cohort study that uses Optum's de-identified Market Clarity Data (2007–2023), which includes integrated medical and pharmacy claims and electronic health record data.
- Data analysis will include descriptive analysis of patient demographics, clinical characteristics and healthcare costs, Kaplan–Meier analysis for time-to-event outcomes (e.g., additional surgeries), and comparison of healthcare costs using statistical tests like Wilcoxon Signed Rank tests.
- The preliminary cohort sizes are 26,472 lumpectomy patients (5-year follow-up), 16,836 mastectomy patients (5-year follow-up) and 6828 NSM, patients of whom, 541 developed nipple necrosis.
- Lumpectomy patients will be followed for prevalence and cost of additional surgeries, subsequent therapies, and healthcare resource utilization. Mastectomy patients will be studied for overtreatment, healthcare costs and long-term outcomes like mortality. Nipple-sparing mastectomy patients will be followed for frequency, distribution and cost of interventions required to resolve nipple necrosis.
- The study leverages a robust dataset reflecting diverse patient and physician practices, linking claims with medical records to provide a comprehensive view of real-world patient journeys and healthcare impacts.
- This study will describe the impact of the suboptimized patient journeys to be from the perspective of multiple stakeholders, including added burden to both the healthcare system and patients, plus financial impact to payors.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: https://becarispublishing.com/doi/epdf/10.57264/cer-2025-0013

Author contributions

SE Wing conducted the original draft preparation. All authors participated in conceptualization, methodology, review and editing.

Financial disclosure

This research was supported by Intuitive Surgical. The authors have received no other financial and/or material support for this research or the creation of this work apart from that disclosed.

Competing interests disclosure

SE Wing, Y Liu and F Zheng are employees of Intuitive Surgical. JC Johnson and KJ Moore are employees of Optum Inc. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing disclosure

No funded writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Institutional review board approval or waiver of approval was not required for this study because the study data were secondary and de-identified in accordance with the US Department of Health and Human Services Privacy Rule's requirements for de-identification codified at 45 C.F.R. § 164.514(b).

References

1. Globocan 2022. *International Agency for Research on Cancer* (2024).
2. Duggan C, Trapani D, Ilbawi AM *et al.* National health system characteristics, breast cancer stage at diagnosis, and breast cancer mortality: a population-based analysis. *Lancet Oncol.* 22(11), 1632–1642 (2021).
3. Nelson JA, Rubenstein RN, Haglich K *et al.* Analysis of a trend reversal in US lumpectomy rates from 2005 through 2017 using 3 nationwide data sets. *JAMA Surg.* 157(8), 702–711 (2022).
4. Rancati A, Gercovich FG. Introduction to conservative mastectomies. *Gland Surg.* 4(6), 450–452 (2015).
5. Wong SM, Chun YS, Sagara Y, Golshan M, Erdmann-Sager J. National patterns of breast reconstruction and nipple-sparing mastectomy for breast cancer, 2005–2015. *Ann. Surg. Oncol.* 26(10), 3194–3203 (2019).
6. Bailey CR, Ogbuagu O, Baltodano PA *et al.* Quality-of-life outcomes improve with nipple-sparing mastectomy and breast reconstruction. *Plastic Reconstr. Surg.* 140(2), 219–226 (2017).
7. Giaquinto AN, Sung H, Miller KD *et al.* Breast cancer statistics, 2022. *CA Cancer J. Clin.* 72(6), 524–541 (2022).
8. Morrow M, Abrahamse P, Hofer TP *et al.* Trends in reoperation after initial lumpectomy for breast cancer: addressing overtreatment in surgical management. *JAMA Oncol.* 3(10), 1352–1357 (2017).
9. Mo C, Ruan W, Lin J, Chen H, Chen X. Repeat breast-conserving surgery versus salvage mastectomy for ipsilateral breast tumour recurrence after breast-conserving surgery in breast cancer patients: a meta-analysis. *Front. Oncol.* 11, 734719 (2021).
10. Sorrentino L, Agozzino M, Albasini S *et al.* Involved margins after lumpectomy for breast cancer: always to be re-excised? *Surg. Oncol.* 30, 141–146 (2019).
11. Kenne Sarenmalm E, Öhlén J, Odén A, Gaston-Johansson F. Experience and predictors of symptoms, distress and health-related quality of life over time in postmenopausal women with recurrent breast cancer. *Psychooncology* 17(5), 497–505 (2008).
12. Pataky RE, Baliski CR. Reoperation costs in attempted breast-conserving surgery: a decision analysis. *Curr. Oncol.* 23(5), 314–321 (2016).
13. Boscoe FP, Johnson CJ, Henry KA *et al.* Geographic proximity to treatment for early stage breast cancer and likelihood of mastectomy. *Breast* 20(4), 324–328 (2011).
14. Landercasper J, Ramirez LD, Borgert AJ *et al.* A reappraisal of the comparative effectiveness of lumpectomy versus mastectomy on breast cancer survival: a propensity score–matched update from the national cancer data base (NCDB). *Clin. Breast Cancer* 19(3), e481–e493 (2019).
15. Hanson SE, Lei X, Roubaud MS *et al.* Long-term quality of life in patients with breast cancer after breast conservation vs mastectomy and reconstruction. *JAMA Surg.* 157(6), e220631–e220631 (2022).
16. Kwok AC, Goodwin IA, Ying J, Agarwal JP. National trends and complication rates after bilateral mastectomy and immediate breast reconstruction from 2005 to 2012. *Am. J. Surg.* 210(3), 512–516 (2015).
17. Smith BD, Jiang J, Shih YC *et al.* Cost and complications of local therapies for early-stage breast cancer. *J. Natl Cancer Inst.* 109(1), djw178 (2017).
18. Caruso F, Ferrara M, Castiglione G *et al.* Nipple sparing subcutaneous mastectomy: sixty-six months follow-up. *Eur. J. Surg. Oncol.* 32(9), 937–940 (2006).
19. Crowe JP, Patrick RJ, Yetman RJ, Djohan R. Nipple-sparing mastectomy update: one hundred forty-nine procedures and clinical outcomes. *Arch. Surg.* 143(11), 1106–1110; discussion 1110 (2008).

20. Garwood ER, Moore D, Ewing C *et al.* Total skin-sparing mastectomy: complications and local recurrence rates in 2 cohorts of patients. *Ann. Surg.* 249(1), 26–32 (2009).
21. Komorowski AL, Zanini V, Regolo L, Carolei A, Wysocki WM, Costa A. Necrotic complications after nipple- and areola-sparing mastectomy. *World J. Surg.* 30(8), 1410–1413 (2006).
22. Mortenson MM, Schneider PD, Khatri VP *et al.* Immediate breast reconstruction after mastectomy increases wound complications: however, initiation of adjuvant chemotherapy is not delayed. *Arch. Surg.* 139(9), 988–991 (2004).
23. Petit JY, Veronesi U, Luini A *et al.* When mastectomy becomes inevitable: the nipple-sparing approach. *Breast* 14(6), 527–531 (2005).
24. Rusby JE, Smith BL, Gui GP. Nipple-sparing mastectomy. *Br. J. Surg.* 97(3), 305–316 (2010).
25. Wagner JL, Fearmonti R, Hunt KK *et al.* Prospective evaluation of the nipple-areola complex sparing mastectomy for risk reduction and for early-stage breast cancer. *Ann. Surg. Oncol.* 19(4), 1137–1144 (2012).
26. National Comprehensive Cancer Network. Breast Cancer (Version 3.2022). (Accessed: 22 May 2025).
27. U.S. Department of Health and Human Services. 45 CFR § 164.514 - Other requirements relating to uses and disclosures of protected health information. (2025). https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-C/part-164/subpart-E/section-164.514
28. U.S. Department of Health & Human Services. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. (2012). https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html
29. Geiger CK, Sheinson D, To TM, Jones D, Bonine NG. Real-world clinical and economic outcomes among persons with multiple sclerosis initiating first- versus second- or later-line treatment with ocrelizumab. *Neurol. Ther.* 12(5), 1709–1728 (2023).
30. Kim LS, Moore MS, Schneider E *et al.* National patterns of hospital admission versus home recovery following mastectomy for breast cancer. *Ann. Surg. Oncol.* 31(13), 9088–9099 (2024).
31. Nattinger AB, Laud PW, Bajorunaite R, Sparapani RA, Freeman JL. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv. Res.* 39(6 Pt 1), 1733–1749 (2004).
32. Centers for Medicare & Medicaid Services. Medicare Severity Diagnosis Related Group (MS-DRG) classifications and software. (2020). Accessed: May 2025: https://www.cms.gov/medicare/coding/medicare-severity-diagnosis-related-group-ms-drg-classifications-and-software
33. Zheng G, Leone JP. Male breast cancer: an updated review of epidemiology, clinicopathology, and treatment. *J. Oncol.* 2022, 1734049 (2022).
34. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med. Care* 36(1), 8–27 (1998).
35. Bayliss EA, Ellis JL, Shoup JA, Zeng C, McQuillan DB, Steiner JF. Association of patient-centered outcomes with patient-reported and ICD-9-based morbidity measures. *Ann. Fam. Med.* 10(2), 126–133 (2012).