Appendix 1 (as supplied by the authors): Cohort dataset creation plan

Study Design

This research ethics board approved, retrospective population based cohort study linked data from population-based data sets available from the Institute for Clinical Evaluative Sciences (ICES) to capture perioperative (i.e. before and after definitive breast cancer surgery) imaging in patients with primary operable (i.e. early stage breast cancer) in Ontario.

Datasets

Patient data was linked from three datasets housed at ICES accessed from the ICES Ottawa site:

Ontario Cancer Registry (OCR): The OCR contains information on all incident cancer diagnoses in Ontario, except non-melanoma skin cancer(1);

Discharge Abstract Database (DAD). The Canadian Institute for Health Information's DAD contains demographic, diagnostic, procedural, and treatment information for all acute care hospitalizations in Canada(2) ;

Ontario Health Insurance Plan (OHIP) database. The OHIP database captures health services billing claims paid by the Ontario Health Insurance Plan to providers. Each row in the OHIP database records the patient, provider, and diagnosis/procedure being claimed for remuneration(2) .

Cohort Creation

STEP 1: Identify primary breast cancer patients from the OCR

The OCR was used to identify all unique adult female breast cancer patients diagnosed with stage I and II breast cancer between January 1, 2007 and December 31, 2012. These dates were chosen because: 1) ICES has access to valid collaborative staging data, containing information on tumour characteristics such as grade and HER-2 status, from 2007 onwards; and 2) 2012 is the last year that OCR data is available at ICES. The OCR determines patient stage by pulling from the electronic medical records after all the patient's tests have been received; as such this stage is considered the 'best' available stage.. Breast cancer diagnosis was based on ICD-9 code 174 – 'malignant neoplasm of female breast'. Adult patients were considered to be those aged greater or equal to 20 years. Patients were excluded if they had a prior breast cancer diagnosis as determined based on OCR variable 'prim' which is the number assigned chronologically by the OCR to a patient's resolved case records based on the diagnosis date. Patients were also excluded if they had stage 0, III, null, or unknown stage disease; Ductal Carcinoma In Situ cases were excluded because ICES does not receive in-situ cases from the OCR. Staging information was based on the OCR 'best stage' variable, which is determined algorithmically based on resolution between the clinical and pathological stage(3) in that the OCR determines patient stage by pulling from the EMRs after all the patient's evaluations have been received; as such this stage is considered the 'best' available stage.. Patients with the following OCR best stage values were included in the model: I, IA, IB, II, IIA, IIB.

STEP 2: Identify primary operable patients by linking OCR with DAD

The OCR cohort was then linked with the DAD to identify patients who underwent breast related surgeries. All procedural codes were based on the Canadian Classification of

Health Interventions (CCI), which is the Canadian standard for classifying health-related

interventions (Canadian Institute for Health Information, 2014).

2a: Identify patients with breast related surgeries

We identified all patients who received any of the following breast related

procedures as their main surgical interventions (DAD variable *incode1*): breast conserving

surgery, mastectomy, axillary lymph node dissection, or sentinel lymph node biopsy. All

CCI intervention codes are shown in Appendix 2.

2b. Identify patients who had surgery within 3 month window after diagnosis

In order to identify primarily operable patients, the study team mandated that only

patients with a first diagnosis of breast cancer who had their definitive surgery within 3

months of their tissue diagnosis date be included in the study population. This window

was selected in an effort to exclude patients who may have received pre-operative systemic

therapy for initially inoperable locally advanced disease as contemporary pre-operative

chemotherapy regimens are a minimum of 4-5 months in duration(4); and these patients

are routinely imaged prior to initiation of induction therapy.

For patients with more than one surgery on the same breast within the three month

window we handled duplicate surgeries in this way: 1) if the diagnoses were the same for

both surgeries then we included the last surgery only, considering this the definitive

surgery on that breast; 2) if the diagnoses were different then we included the first surgery

only, as this corresponded to the first diagnosis of breast cancer.

Breast cancer patients were excluded if they did not have any breast or axillary surgeries, if their surgery date was on or before their cancer diagnosis date, or if their surgery date was more than three months after their diagnosis date; the cohort creation flow is shown in Figure 1 (main article).

STEP 3: Identify covariates

To identify factors associated with utilization of imaging patient level data was collected for: age at diagnosis, disease stage (I, II, III), histology (characterized as ductal, lobular, or other), comorbidity calculated using the Deyo modification of the Charlson comorbidity index(5) (characterized as 0, 1-2, 3+), surgery type (breast conserving surgery vs. non breast conserving surgery), whether additional preoperative loco-regional imaging beyond mammography was performed (either breast ultrasound or breast MRI), and neighbourhood income quintile based on Postal Code Conversion File Plus(6). Information regarding institution type, patient rurality at diagnosis (based on Rurality Index Ontario 2008(7)), and Local Health Integration Network (LHIN) of the hospital performing the surgery was also collected. Ontario is divided geographically into 14 LHINs which administer and coordinate local health systems(8).

The OCR does not include all the clinical information required for treatment decisions so we used the OCR's Collaborative Staging dataset to obtain additional disease characteristics not available in the OCR. This included; tumor grade (1, 2, 3), HER-2 status (positive or negative), triple negative disease (positive or negative), axillary lymph node involvement (yes or no) and presence of lymphovascular invasion (yes or no). For determining HER-2 status if a

Fluorescence In Situ Hybridization (FISH) test was performed this value was used, otherwise we used the imunohistchemistry test with all borderline/indeterminate results classified as negative.

Imaging quantification and classification:

To quantify all perioperative staging imaging the cohort of primary operable breast cancer patients was linked with Ontario Health Insurance Plan (OHIP) data to obtain all billing fee codes charged for these patients between January 1, 2007 and December 31, 2012. Subjects were followed until 3 months after their index surgery event; the maximum follow up date was March 31, 2013. All OHIP fee codes were reviewed by DS and those not considered relevant to breast cancer staging imaging were discarded, this resulted in a final list of 126 OHIP fee codes which DS classified by imaging body site and modality (see Appendix 2). A flat file including all imaging fee codes and associated service dates for each patient was created; fee codes only contributed to the final imaging counts if the service date occurred during the perioperative period. During this period, imaging was considered pre-operative if it occurred between the tissue diagnosis date (usually breast biopsy) and the day of surgery and post-operative if it occurred within the 3 month window after the definitive breast surgical procedure. All imaging tests were quantified by site (skeleton, thorax, abdomen and/or pelvis, other) and modality (plain x-rays, ultrasonography, computerized tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) etc.).

Imaging tests were further classified as i) initial imaging or ii) confirmatory imaging. Initial imaging was defined as the first imaging test performed on a body site (i.e. skeleton, abdomen, or thorax) in the pre-specified time period. Confirmatory imaging was defined as any additional imaging test performed on a body site already imaged. We did not count any other

imaging tests of the same anatomical body area beyond the first confirmatory scan in the total

imaging counts. This conservative approach was decided by the team to try and exclude any

further imaging that could potentially be unrelated to the purposes of breast cancer staging.

To determine who ordered the imaging test we extracted the referring physician's self-

reported main specialty associated with the OHIP imaging fee code. Physicians were classified

as: surgery, medical oncology (including haematology and gynaecological oncology), radiation

oncology, primary care (including internal medicine), diagnostic imaging (including nuclear

medicine), and other.

Statistical analysis

The total number of primary and confirmatory imaging tests for stages I and II, combined

and individually were calculated, considering both the number of patients imaged and the

number of imaging tests performed. Initial and confirmatory imaging was further classified into

pre- and post-operative imaging. All initial and confirmatory imaging tests were quantified by

body site (skeleton, thorax, abdomen and/or pelvis, other) and imaging modality. To assess

whether use of imaging modalities was changing over time the percentage of patients imaged

each year was calculated by body site and over all body sites and considering initial and

confirmatory separately. Trends were evaluated using the Cochran-Armitage test.

To identify characteristics associated with metastatic imaging utilization we classified

patients categorically into five imaging groups based on whether they received imaging (yes or

no) on: i) any site (skeleton, thorax, abdomen and/or pelvis, other), ii) skeleton, iii) thorax, iv)

abdomen and/or pelvis, or v) other). Metastatic imaging did not include any imaging performed

on the breast, such as US, MRI or mammogram. For skeleton imaging we did not count skeleton

MRIs that were billed on the same day as a breast MRI, the rational being that if breast MRI and skeleton MRI were billed concurrently on the same day then the skeleton MRI was just part of the breast MRI and did not qualify as further staging imaging. For each category, multivariate logistic regression was performed to identify factors associated with imaging use performed per body site; this analysis was based on the subset of patients for whom Collaborative Staging variables were available. Predictors were included as categorical variables and results were reported as odds ratios. All tests used $p = 0.05$ (two-sided) to evaluate statistical significant. Analyses were performed using SAS software, version 9.3 (SAS Institute).

REFERENCES

1.   Ontario CC. Health System Data 2013. Available from: https://www.cancercare.on.ca/toolbox/systeminfo/.
2.   Care OMoHaL-T. Health Analyst's Toolkit 2012. Available from: http://www.health.gov.on.ca/english/providers/pub/healthanalytics/health_toolkit/health_toolkit.pdf.
3.   Walters S, Maringe C, Butler J, Brierley JD, Rachet B, Coleman MP. Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership. International journal of cancer Journal international du cancer. 2013 Feb 1;132(3):676-85. PubMed PMID: 22623157.
4.   Untch M, Konecny GE, Paepke S, von Minckwitz G. Current and future role of neoadjuvant therapy for breast cancer. Breast. 2014 Oct;23(5):526-37. PubMed PMID: 25034931.
5.   Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. Journal of clinical epidemiology. 1992 Jun;45(6):613-9. PubMed PMID: 1607900.
6.   Canada S. Postal Code Conversion File Plus (PCCF+) (82F0086X) 2014 [cited 2014 December 20, 2014]. Available from: http://www5.statcan.gc.ca/olc-cel/olc.action?objId=82F0086X&objType=2&lang=en&limit=0.
7.   OMA Economics Department. Measuring Rurality - RIO2008_BASIC: Methodology and Results 2009. Available from: https://www.oma.org/Resources/Documents/2008RIO-FullTechnicalPaper.pdf
8.   (2006) OMoHaL-TC. Local Health System Integration Act, 2006 2006. Available from: http://web.archive.org/web/20120315221643/http://www.health.gov.on.ca/english/public/legislation/lhins/lhins_faq.html#1.