

UCSF

UC San Francisco Previously Published Works

Title

Linkage of the ACR National Mammography Database to the Network of State Cancer Registries: Proof of Concept Evaluation by the ACR National Mammography Database Committee

Permalink

<https://escholarship.org/uc/item/3bf1f68t>

Journal

Journal of the American College of Radiology, 16(1)

ISSN

1546-1440

Authors

Zuley, Margarita L
Nishikawa, Robert M
Lee, Cindy S
et al.

Publication Date

2019

DOI

10.1016/j.jacr.2018.06.027

Peer reviewed



Published in final edited form as:

J Am Coll Radiol. 2019 January ; 16(1): 8–14. doi:10.1016/j.jacr.2018.06.027.

Linkage of the ACR National Mammography Database to the Network of State Cancer Registries: Proof of Concept Evaluation by the ACR National Mammography Database Committee

Margarita L. Zuley, MD^a, Robert M. Nishikawa, PhD^b, Cindy S. Lee, MD^c, Elizabeth Burnside, MD^d, Robert Rosenberg, MD^e, Edward A. Sickles, MD^f, Wendie Berg, MD, PhD^a, Jessica Leung, MD^g, Jennifer Harvey, MD^h, Debapriya Sengupta, MBBS, MPHⁱ, and David Gur, ScD^b

^aDepartment of Radiology, Division of Breast Imaging, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

^bDepartment of Radiology, Imaging Research Division, University of Pittsburgh, Pittsburgh, Pennsylvania.

^cDepartment of Radiology, New York University Langone Health, Garden City, New York.

^dDepartment of Radiology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin.

^eRadiology Associates of Albuquerque PA, Albuquerque, New Mexico.

^fDepartment of Radiology and Biomedical Imaging, University of California, San Francisco, California.

^gDiagnostic Radiology Department, The University of Texas MD Anderson Cancer Center, Houston, Texas.

^hDepartment of Radiology, University of Virginia Health System, Charlottesville, Virginia.

ⁱQuality and Safety, American College of Radiology, Reston, Virginia.

Abstract

Purpose: The National Mammography Database (NMD) contains nearly 20 million examinations from 693 facilities; it is the largest information source for use and effectiveness of breast imaging in the United States. NMD collects demographic, imaging, interpretation, biopsy, and basic pathology results, enabling facility and physician comparison for quality improvement. However, NMD lacks treatment and clinical outcomes data. The network of state cancer registries (CRs) contains detailed pathologic, treatment, and clinical outcomes data. This pilot study assessed electronic linkage of NMD and CR data at a multicenter institution as proof of concept.

Corresponding author and reprints: Margarita Zuley, MD, Department of Radiology, Division of Breast Imaging, University of Pittsburgh Medical Center, Magee Womens Hospital, 300 Halket Street, Pittsburgh PA 15213; zuleyml@upmc.edu.

ADDITIONAL RESOURCES

Additional resources can be found online at: <https://doi.org/10.1016/j.jacr.2018.06.027>.

Materials and Methods: We obtained Quality Oversight Committee approval for this retrospective study. Data of patients diagnosed with breast cancer in 2014 and 2015 were retrieved from our NMD-approved radiology information system (RIS) and matched with reportable patients in our CR using social security number (SSN), first name (fname), last name (lname), and date of birth (DOB). Matching was repeated without SSN. Percentage and reasons for mismatch were evaluated.

Results: The RIS query identified 1,316 patients. CR linkage was 99.2% successful (n = 1,305 of 1,316) using SSN, fname, lname, and DOB. Eleven mismatches included four CR case-finding failures, one NMD fname error, five nonreportable in the CR, and one with correct identifiers in both databases. Without SSN, linkage was 97.3% successful (n = 1,281 of 1,316); name errors accounted for 19 and DOB accounted for 5 additional mismatches.

Conclusion: Using common data elements, linkage between the NMD and state CRs may be feasible and could provide critical outcomes information to advance accurate assessment of breast imaging in the United States.

Keywords

Breast; data linkage; record linkage; medical

INTRODUCTION

The ACR National Mammography Database (NMD) is part of the National Radiology Data Registry (NRDR), which is overseen ultimately by the ACR Commission on Quality and Safety. The NMD Committee and researchers at the University of Pittsburgh present a proof of concept electronic linkage of the NMD and cancer registries (CRs) data at a multicenter institution.

The NMD began in 2008 and has become a powerful auditing aid for participating facilities across the United States. The NMD currently contains data for more than 19,500,000 mammograms, the large majority of which are current digital technology, making it the largest and potentially most comprehensive source of information regarding the use and effectiveness of breast imaging in the United States (Judy Burleson, ACR, personal communication November 6, 2017). The NMD collects detailed information on patient demographics, imaging examinations performed, BI-RADS assessments, and percutaneous biopsy results and includes optional fields for staging [1]. The NMD currently has 693 participating facilities from 39 US states and territories. The NMD allows facilities to compare their individual physician performances, as well as overall group performance levels, against other participants and against published benchmarks. Each participating facility must have an NMD-certified vendor or ACR-approved home-grown system that uploads data to the ACR registry using a standard query created by ACR for NMD-defined common data elements within a defined time frame. However, the NMD suffers from reliable entry of surgical pathology data, because often neither the facility radiology information system (RIS) nor the NMD registry is linked to pathology databases, thus requiring manual entry of information by the facility staff. As a result, many participating facilities do not include detailed surgical pathology data, and those that do have expected

limitations related to human error during data entry. This lack of surgical pathology data may be particularly important for cases in which the percutaneous image-guided breast biopsy is benign or has high-risk histology, and the subsequent surgical biopsy does demonstrate cancer. Also, the NMD suffers from the lack of important clinical outcomes data. The linking of the NMD to state CRs that collect comprehensive demographic, pathologic, treatment, and mortality information has the potential to significantly improve the completeness and overall quality of outcome measures for breast imaging. Such linkage would, for example, provide robust data to answer current controversies surrounding the relative roles of screening frequency, ages at first and last screen, and screening history versus treatment regimen for mortality and morbidity. Such linkage would also allow a clear understanding of how various screening paradigms affect interval cancers types and rates.

In 1992, President Clinton signed the Cancer Registries Amendment Act, which requires every state to maintain a CR. The Centers for Disease Control and Prevention (CDC) has been charged with administering this act, which it does via the National Program of Cancer Registries (NPCR). Annually, each participating US state or territory submits its registry data to the NPCR using data originating from medical facilities, such as hospitals or medical practices. The local data are entered by certified registrars who adhere to strict rules set forth by the North American Association of Central Cancer Registries (NAACCR), an organization that certifies registries, develops and promotes standards for data entry, educates and trains registry professionals, and aggregates and publishes central data. Currently through the NPCR, the CDC supports central CRs in 46 states, the District of Columbia, Puerto Rico, the US Pacific Island Jurisdictions (including Guam, American Samoa, and the Commonwealth of the Northern Mariana Islands), and the US Virgin Islands. The NPCR represents 97% of the US population. Arizona, New Mexico, Utah, and South Dakota do not participate in the NPCR program. However, Arizona and South Dakota maintain the mandated registries for their states, and Utah and New Mexico participate in the Surveillance, Epidemiology, and End Results (SEER) registries. As of July 2017, 44 of the 52 NPCR registries have signed the NAACCR National Interstate Exchange Agreement, which enables information sharing with the other registries, thus improving continuity of the patient record [2]. The case-finding process used by every facility's CR is a carefully defined and quality-controlled process intended to identify every patient diagnosed or treated with a reportable diagnosis. According to the NPCR guidelines, all cases with International Classification of Diseases for Oncology behavior codes of 2 (in situ) and 3 (malignant, primary site) must be included in the registry. In addition, the facility cancer committee or their state central registry may also require benign or high-risk cases to be included. Along with national, state, and local requirements, CRs participating in the Approvals Program of the Commission on Cancer of the American College of Surgeons must use the reportable list defined by the combination of requirements from all these authorities.

The other authoritative source of cancer outcomes data is the National Cancer Institute (NCI)-funded SEER program. This program started in 1973 and has expanded to currently collect information from 10 states (6 of which also are NPCR-participating states) and 10 geographic regions (Table 1). The SEER registries represent 28% of the US population and include high percentages of minority populations [3] (Fig. 1). A recent evaluation of cancer mortality from 1980 to 2014 by smaller areas across the country demonstrated a large range

of mortality reduction for breast cancer. In part, the authors attributed the differences in mortality across the nation to variable screening access and appropriate therapy [4]. Thus, outcomes analysis based purely on SEER registries may not be reflective of the nation because SEER data only include a relatively small fraction of the population.

In the past, the largest source of outcomes data for mammography in the United States has been the NCI-funded Breast Cancer Surveillance Consortium (BCSC). NCI established the BCSC in 1994 because of a mandate from the 1992 Mammography Quality Standards Act. This research consortium of seven state or geographic SEER registries was intended to sample approximately 2% of the US population to demonstrate effectiveness of screening. Because data on mammography screening in underserved populations was a priority of NCI, consortium members specifically included those applicants that included rural and minority populations [5]. Although an excellent source of data for many years, BCSC has suffered in the more recent past with diminished funding, which has led to closures of two of the seven registries. Because of its age, the BCSC data are primarily based on screen-film mammography, a technology that is now outdated and replaced with digital imaging. Furthermore, 35.3% of women within the BCSC had only one mammogram available in the database, and an additional 17.9% had data for only two mammograms entered. Thus, less than 50% of women included in this data set have had more than two mammography results recorded for analysis. As such, robust conclusions regarding the relationship of current technology screening mammography to clinical outcomes based on the BCSC data are questionable.

Currently, controversy exists as to the benefits versus risks of screening, and clear assessment of screening performance is not possible in this country, in part because cancer outcomes data are not linked directly to screening data for most of the population. Such linkage would substantially inform decision making by enabling robust analysis of important currently debated topics, such as effects of screening interval, onset and termination of screening, types of surveillance (such as supplemental screening ultrasound, tomosynthesis, or MRI), and impact of different therapies by patient, tumor, and treatment details. Linking the NMD to the CDC network of CRs would address this deficiency. The purpose of this study was to attempt electronic linkage between the NMD data and CR data at a multicenter institution as a proof of concept.

MATERIALS AND METHODS

NMD Extraction

This retrospective study was approved by our enterprise Quality Oversight Committee. Query of our NMD 2.0-approved RIS named “MAM NMD Download Report” (Imagecast, General Electric, Milwaukee, Wisconsin), which we routinely run and upload to ACR semi-annually as part of our quality program, was used for this study. The RIS collects demographic information from our hospital billing systems (Medipac [Allscripts, Chicago, Illinois] and EPIC [Verona, Wisconsin]) and the results of the imaging and percutaneous biopsy from manual data entry by trained breast imaging staff. Results of the manual data entry are routinely audited for completeness and accuracy through a carefully defined

institution-wide auditing process. All NMD-defined data elements collected from the standard ACR-created query of the RIS are listed in Appendix 1.

In January 2017, the archived MAM NMD Download Report queries for 2014 and 2015 data that had been run and uploaded to ACR as part of our normal quality process were accessed for this study. Eighteen separate archived reports were used because the queries were made separately for each of our nine breast imaging facilities for two date ranges: from January 1, 2014, 00:00 until December 31, 2014, 23:59 and January 1, 2015, until December 31, 2015, 23:59. These reports exist as .txt files for each facility for each date range. These 18 facility-level reports were then copied into a single Excel (Microsoft, Redmond, Washington) file for our project. This file included every encounter of every patient seen in any of our breast imaging facilities from January 1, 2014, to December 31, 2015. The NMD data element titles were manually added to the research Excel file as column headers, because the clinical NMD download report does not include headers, and then the file was sorted for the column "classification of lesion" having a code of 3 (which indicates malignancy). All lines in both files with a classification of lesion = 3 were exported to a new Excel file. This final file, named MAL-3, was used for matching with the CR.

CR Extraction

Simultaneously, in January 2017, a query of our CR, METRIQ (Elekta Medical Systems, Stockholm Sweden), was made for all reportable cases with a diagnosis of breast cancer from January 1, 2014, to January 1, 2017. The date range for the CR search was extended beyond the 2-year study window to allow inclusion of patients who were diagnosed in 2014 or 2015 but did not receive treatment until after the end of 2015. Reportable cases are defined by the CDC and state-specific central CRs. For our state, reportable cases include all patients whose initial surgery was performed within our institutional network, as well as all patients who received additional treatment in one of our facilities after initial diagnosis or surgery. (The CR also includes nonreportable cases. Nonreportable cases include, among other things, all patients who had a diagnosis of breast cancer by imaging directed biopsy but did not undergo surgery or additional treatment for that cancer at any of our institutional facilities.)

The sources of case-finding for our network of registries are multiple and redundant. First, the hospital billing systems (Medipac and EPIC) use the International Classification of Diseases, 10th revision codes for every patient episode. Second, our pathology database (CoPath [Cerner Corporation, Kansas City, Missouri]) is used to ascertain all neoplastic specimens assigned a Systematized Nomenclature of Medicine code into the range of 80002 to 9989. Third, the radiation oncology electronic medical record (ARIA [Varian Medical Systems, Palo Alto, California]) identifies all patients treated with radiation. All cases identified from any of these sources are then manually reviewed by qualified CR staff and abstracted into the registry. The manual abstraction process is routinely audited for accuracy and completeness using a standardized process. The list of all data elements from the CR included in the trial are listed in Appendix 1.

The CR query executed for this study in January 2017 was accomplished by searching METRIQ for all cases abstracted into any hospital database within our network of CRs from

January 1, 2014, to January 2017 with data field “primary site” containing C500 to C509 (meaning primary breast malignancy). The results of the query were exported to an Excel file. The two research Excel files (“MAL-3” final RIS research file containing the NMD data and the reportable breast cancer case file from the CRs) were then reviewed, and differences in common data element column configurations were corrected. These corrections included removal of erroneous spaces in column headers, invalid characters, and extra commas; reformatting the social security number (SSN) column in the MAL-3 file to remove dashes; and the addition of leading “0” for SSN column in the MAL-3 file, which were eliminated during the copy and paste into the final Excel file because the column was formatted as “general” as opposed to “text.” After these corrections were complete, the two files were imported into our SQL server (Microsoft) and electronically matched using the LEFT JOIN key word and the criteria of first name, last name, date of birth (DOB), and SSN in both files. The LEFT JOIN key word returns all records from the left table (in this study the MAL-3 file) and only the matched records from the right table (the CR file). All records not found in the right file are listed as NULL, thus cases listed as NULL in the CR side of the match were the unmatched cases. The resulting merged file was saved. The electronic match and save was repeated without SSN. This resulted in two final files. Unmatched cases, which contained NULL in the CR columns, were analyzed for reason of matching failure by manual review of each failed instance (as depicted in Fig. 2).

RESULTS

The archived RIS queries revealed 164,827 lines representing 121,326 patients based on unique SSN and 43,501 duplicate lines. Duplicates occurred because some patients were seen more than one time in the 2-year study period. Sorting of the file by the column “classification of lesion” revealed 1,900 lines having a value of 3 (meaning malignancy), including 1,316 patients based on SSN and 584 duplicates. Reasons for duplication included 577 patients who underwent screening initially, were recalled for diagnostic evaluation, and were diagnosed with breast cancer within the 2 study years (true-positive examinations); 6 patients who developed an interval breast cancer after a benign interpretation in the 2-year period (false-negative examinations); and 1 patient who had transposition of two digits of her SSN during one of her visits due to a data entry error.

Search of our institutional network of CRs identified 4,158 lines with CR data element “primary site” starting with C50 (which indicates breast primary). Reasons for duplications included patients with multiple primary cancers diagnosed and patients receiving a portion of their care at more than one of our facilities and thus entered into more than one hospital registry. Also, because the search was run from January 1, 2014, to January 1, 2017, patients diagnosed with breast cancer after the 2-year study range were included.

Electronic matching of the extracted files using first name, last name, DOB, and SSN was 99.2% successful (1,305 of 1,316 patients). Eleven cases failed to electronically match, including 10 of 1,316 (0.7%) patients on the NMD list as having a new breast cancer diagnosis but not entered into the CRs database, and 1 patient(0.1%) who did not electronically match despite being included and having all personal health identifiers match in both database lists. Of the 10 cases missing from the CR, 4 were due to failure of the CR

abstraction process (which is a manual chart review performed by certified personnel), 5 had a percutaneous biopsy at one of our breast imaging facilities and were entered into our RIS but did not receive treatment at any of our hospitals and therefore were nonreportable, and 1 patient had a first name mismatch (entered as Mary in one database and Margaret in the other). Repeat matching without SSN using only first name, last name, and DOB revealed 24 additional unmatched patients, including 14 patients with a mismatch in last name (apostrophes, maiden versus married last name, hyphenated names), 5 patients with mismatch of first name (nickname versus full first name or some misspellings), and 5 patients with an error in their DOB with either the day or year off by one digit.

DISCUSSION

Linkage of imaging history to CRs will allow a much more direct and robust understanding of the role that imaging plays in outcomes from breast cancer. This proof of concept trial has demonstrated that linkage is possible and can be highly successful without substantial manual intervention. With more than 99% of patients successfully paired between the two databases when SSN was used and more than 97% pairing without SSN, the ability to perform this matching on a much larger scale seems to be feasible. In the United States, every state is required by federal law to maintain a certified CR that adheres to at least NAACCR standards of data entry.

There is a vast amount of information available to understand outcomes of breast cancer based on patient demographics, stage at diagnosis, and treatment provided. Linkage of this rich information source to imaging data would provide a key missing element that would help to resolve the ongoing debate regarding the role of imaging, not only in affecting mortality but also potentially the impact of screening on other end points, such as need for various expensive and sometimes toxic treatments for more advanced disease. In addition, interval cancer rates and potential variable interval rates by tumor biology would be able to be clearly identified and analyzed. This comprehensive picture could inform appropriate imaging paradigms by identifying when to start and end standard screening, what modalities have the best accuracy for detecting breast cancers that are clinically important, and what imaging is not beneficial to patient outcomes. Current debate regarding the frequency and the biology of interval cancers in screened patients, the role of improved treatment versus screening in overall and disease-free survival, and overdiagnosis may be directly answered with actual patient outcomes rather than extrapolated data that currently exist (namely, estimates from statistical models, extrapolation from results of smaller cohorts, and outcomes from other countries, most of which have different population risk profiles, screening intervals, recall thresholds, or treatment paradigms from the United States).

The BCSC has provided some insight into these questions, but the data have not been without limitations. From 1994 to 2009, the BCSC collected data for approximately 9.5 million mammograms performed on 2,300,000 women from up to seven registries, including the San Francisco Registry, the Colorado Mammography Advocacy Project (which closed in 2006 because of loss of funding), the New Mexico Mammography Project (which closed in 2010 because of loss of funding), the Metro Chicago Cancer Registry, the Vermont Breast Cancer Surveillance System, the New Hampshire Mammography Network, the Kaiser

Permanente WA Registry, and the Carolina Mammography Registry. From 2010 to 2016, the NCI's Healthcare Delivery Research Program funded a contract with the BCSC to establish it as a resource for external researchers. Recent publications intended to address the role that imaging plays in outcomes based on these registries have inferred end points based on statistical models, not actual patient data [6–8]. Because these registries represent less than 2% of the United States population, have data on three or more screening examinations for less than 50% of the patients, and are primarily based on outdated screen-film mammography, conclusions based on these data as to the effectiveness of screening in altering mortality in the United State are questionable. By linking the ACR NMD registry, which currently contains more than 19,500,000 examinations (the vast majority of which are current technology digital mammograms), to the wide network of CDC-managed state CRs, more direct evaluations may be possible. One limitation of NMD, however, is that participation is voluntary and so there is the possibility that despite the database being larger, it also is not entirely reflective of the population. Currently, fewer than 700 facilities participate in the NMD, which account for 7.9% of the 8,726 facilities accredited to providing mammography services in the United States, per the FDA [9]. Thus, facilities that participate in this voluntary program may not have performance levels reflective of the entire nation.

An unexpected result occurred during this pilot study. We had not considered that the linkage would reveal case-finding errors arising from our CR. Ten patients (0.7%) were not identified by the standard registry case-finding process despite its several redundancies. Our linkage improved the completeness and accuracy of both databases by not only identifying more patients but also identifying errors, such as hyphenated last names, misspellings, and incorrect entries of DOB, as examples.

This study has several limitations. It is a single-institution trial, and therefore the linkage result may reflect the internal quality standards of data capture for our institution. This level of matching may not be attainable at all facilities. Because our local population is not heavily weighted toward an ethnic segment of the population, we did not have a very large population with the same names. In some areas of the United States, this issue may require identifying other patient-level variables for accurate matching. Because patients are not required to provide SSNs any longer, the higher matching achieved in this pilot study may overestimate the success of our linkage. Another limitation is that we performed and analyzed the match by looking for patients in our RIS diagnosed with breast cancer who matched to the CR. Future trials should include matching of CR patients with a primary breast cancer to all patients in the NMD. This would allow near-complete information on false-negative mammograms identified at other institutions, thus creating an important opportunity to discover and learn from false-negative imaging examinations. This type of linkage would be closer to the BCSC methodology, which links all mammograms to CRs. Linking entire mammography registries may create many possible true and false matches of negative mammograms to cancers, requiring more resources to resolve. Some BCSC registries are also linked to benign, as well as malignant, pathology resources. Because some CRs also include patients who underwent benign surgery, this could be an additional avenue for linkage potential. As a next step, we plan to attempt linkage of several other NMD-participating institutions in our region to our state CR to evaluate a larger population. This

process in theory may increase the percentage of patients that link, because it would identify patients diagnosed in one facility but treated in another.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Sharon Winters, MS, CTR, Brenda Crocker, Tom Matus, Divya Bollineni, Mythreyi Bhargavan-Chatfield, PhD, Judy Burleson, MHSA, Lu Myer, Durwin Logue, Amy Klym.

Dr Harvey reports other from Hologic Inc, other from Volpara Solutions, LLC, outside the submitted work. Dr Nishikawa reports grants and other from Hologic, Inc, other from iCAD, Inc, outside the submitted work. Dr Leung reports personal fees from Hologic, Inc, personal fees from Fujifilm, Inc, outside the submitted work. The other authors have no conflicts of interest related to the material discussed in this article.

REFERENCES

1. Lee CS, Bhargavan-Chatfield M, Burnside ES, Nagy P, Sickles EA. The National Mammography Database: preliminary data. *AJR Am J Roentgenol* 2016;206:883–90. [PubMed: 26866649]
2. North American Association of Central Cancer Registries (NAACCR). Available at: <https://www.naacr.org/certified-registries/>. Accessed October 11, 2017.
3. National Cancer Institute. Surveillance, Epidemiology, and End Results Program. Available at: www.seer.cancer.gov. Accessed October 11, 2017.
4. Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, et al. Trends and patterns of disparities in cancer mortality among US counties, 1980–2014. *JAMA* 2017;317:388–406. [PubMed: 28118455]
5. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast cancer surveillance consortium: a national mammography screening and outcomes database. *AJR Am J Roentgenol* 1997;169:1001–6. [PubMed: 9308451]
6. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, et al. Breast Cancer Working Group of the Cancer Intervention and Surveillance Modeling Network. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. *Ann Intern Med* 2009;151:738–47. [PubMed: 19920274]
7. Mandelblatt JS, Stout NK, Schechter CB, van den Broek JJ, et al. Collaborative modeling of the benefits and harms associated with different US breast cancer screening strategies. *Ann Intern Med* 2016;164:215–25. [PubMed: 26756606]
8. Stout NK, Lee SJ, Schechter CB, Kerlikowske K, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. *J Natl Cancer Inst* 2014;106:1–8.
9. US Department of Health and Human Services. US Food & Drug Administration. Available at: <https://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityScorecard/>. Accessed December 27, 2017.

TAKE-HOME POINTS

- Electronic linkage of NMD-approved imaging databases to the network of CRs is feasible and was highly successful in this single-institution pilot study.
- If successful on a larger scale, linkage of the ACR National Mammography Database to the CDC network of state CRs has the potential to address many issues in the long-standing debate regarding the role that imaging plays in outcomes from breast cancer.
- Similar linkage may also be feasible for other registries, such as the ACR lung cancer screening and CT colonography registries.

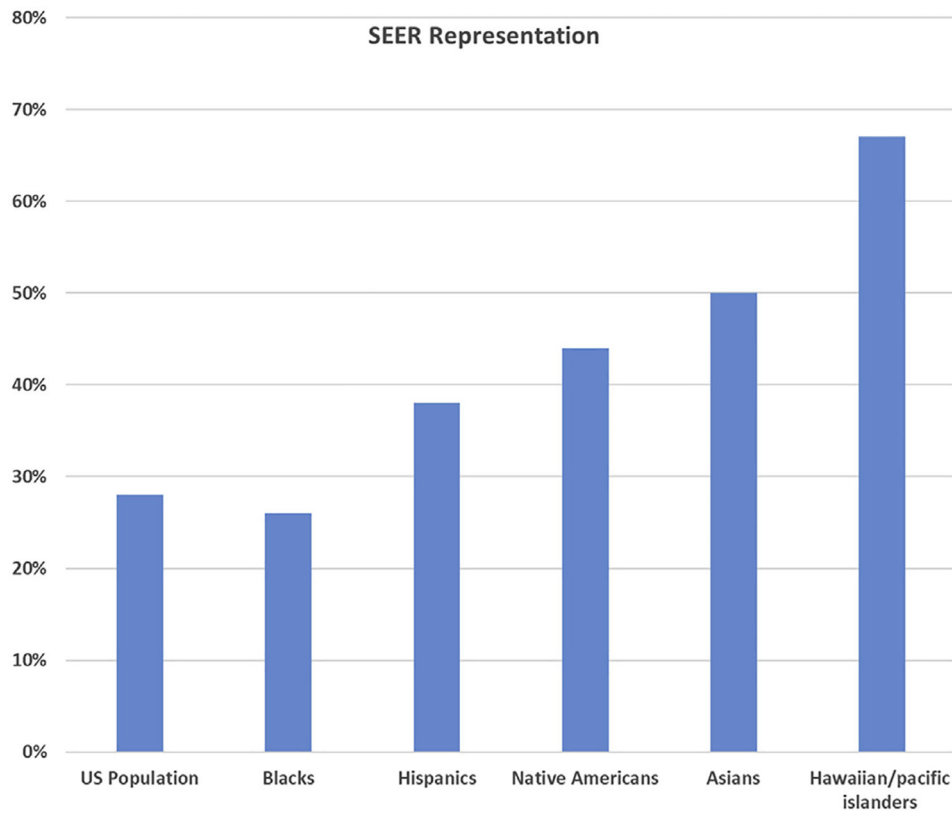


Fig 1. Percentage by ethnicity in United States represented in Surveillance, Epidemiology, and End Results (SEER) registry.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

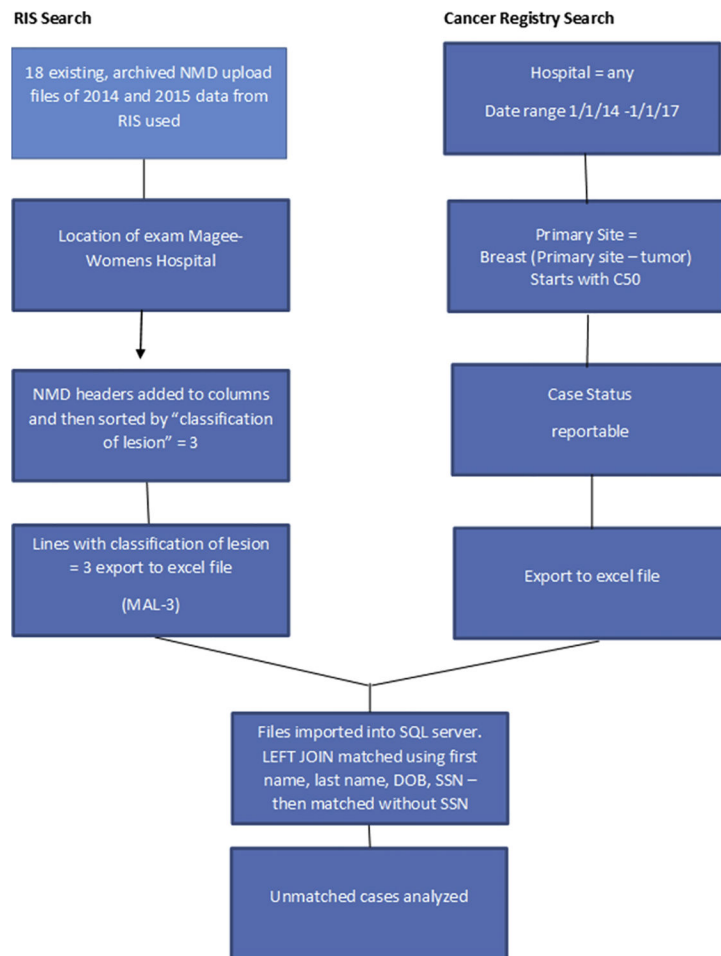


Fig 2. Flowchart of case identification and matching. DOB = date of birth; MAL-3 = final file used for matching with the cancer registry; NMD = National Mammography Database; RIS = radiology information system; SSN = social security number.

Table 1.

States, regions, provinces, and territories participating in Surveillance, Epidemiology, and End Results registry

States	Geographic Regions
Connecticut	Alaska Native Americans
Utah	Arizona Native Americans
Hawaii	Cherokee Nation
Iowa	Detroit
Kentucky	Atlanta
Louisiana	Rural Georgia
New Jersey	Greater Georgia
New Mexico	San Francisco—Oakland
	San Jose—Monterey
	Los Angeles
	Greater California
	Seattle—Puget Sound

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript