Validity of state cancer registry treatment information for adolescent and young adult women

Chelsea Anderson, PhD[1],[*], **Christopher D. Baggett, PhD**[1], **Chandrika Rao, PhD**[2], **Lisa Moy, MPH**[3], **Lawrence H. Kushi, PhD**[3], **Chun R. Chao, PhD**[4], **Hazel B. Nichols, PhD**[1]

[1]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, NC

[2]North Carolina Central Cancer Registry, Raleigh, NC

[3]Division of Research, Kaiser Permanente Northern California, Oakland, CA

[4]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA

Abstract

Background: Population-based cancer registries collect information on first course of treatment that may be utilized in research on cancer care quality, yet few studies have investigated the validity of this information. We examined the accuracy and completeness of registry-based treatment information in a cohort of adolescent and young adult women.

Methods: Women diagnosed with breast cancer, lymphoma, thyroid cancer, cervical/uterine cancer or ovarian cancer at ages 15-39 during 2003-2014 were identified using data from the North Carolina Central Cancer Registry (CCR) (N=2,342). CCR data were linked to Medicaid and private insurance claims data, and claims were reviewed for the 12 months following diagnosis to identify cancer treatments received. Using claims data as the gold standard, we calculated the sensitivity and positive predictive value (PPV) of CCR data for receipt of chemotherapy, radiation and hormone therapy. We also compared dates of treatment initiation between the two data sources.

Corresponding Author: Chelsea Anderson, American Cancer Society, 250 Williams St NW, Atlanta, GA 30303, Chelsea.anderson@cancer.org.

Results: For all cancer types combined, the sensitivity of the CCR data was high for chemotherapy (86%) and moderate for radiation (74%). PPVs were 82% and 83% for chemotherapy and radiation, respectively. Both the sensitivity (67%) and PPV (70%) were lower for hormone therapy for breast cancer. For all three treatment types, dates of initiation in the registry and the claims differed by 30 days for most women.

Conclusions: In this cohort of young women, population-based cancer registry data on chemotherapy receipt was reasonably accurate and complete in comparison with insurance claims. Radiation and hormone therapy appeared to be less complete.

Introduction

Population-based cancer registries are the primary source of cancer incidence and survival statistics in the United States. However, cancer registries also collect information on initial treatment for newly diagnosed cases, and may therefore serve as an important resource for examining cancer care quality, trends in treatment utilization, and disparities in therapy receipt, with implications for efforts to improve outcomes among cancer patients and survivors. Studies utilizing registry-based treatment information for these objectives rely on the assumption that this information is complete and accurate. Yet limited research to date has evaluated the quality of registry-based treatment information in comparison with other data sources such as insurance claims.

Using Medicare claims as the gold standard, a prior report examined the accuracy and completeness of first course of treatment data from the Surveillance, Epidemiology, and End Results (SEER) program, a system of population-based cancer registries across the U.S, among patients aged 65 and older diagnosed with one of several cancer types from 2000 to 2006. They reported that the validity of SEER data was moderate, with sensitivities of 68%, 80%, and 69% for chemotherapy, radiation, and hormone therapy, respectively, and concluded that treatment data from SEER registries should not generally be used for estimating the proportion of individuals treated with these therapies.[1] However, the validity of cancer treatment data from non-SEER registries has seldom been examined, and prior studies have often been limited to breast or prostate cancer.[2–4] Furthermore, to our knowledge, no study has focused on treatment data quality among younger patients who are not generally enrolled in Medicare.

The objective of this study was to assess the validity of first course of treatment information from the North Carolina Central Cancer Registry among adolescent and young adult (AYA, age 15-39 years) women diagnosed with lymphoma, breast cancer, thyroid cancer, or gynecologic cancers, using Medicaid and private insurance claims as the gold standard.

Methods

Data source and study sample

As part of an ongoing study of reproductive outcomes among female AYA cancer survivors, we identified women diagnosed with cancer in North Carolina between 2003 and 2014. Data came from the University of North Carolina Cancer Information and Population Health Resource (CIPHR), which links records from the North Carolina Central Cancer Registry (CCR) to administrative and claims data (through 2015) from private health insurance plans and Medicaid.[5]. The CCR has met the Gold Standard for Registry Certification from the North American Association of Central Cancer Registries (NAACCR) each year since 2008. Either Gold or Silver Certification was achieved each year for 2003-2007. NAACCR certification levels reflect the completeness, accuracy, and timeliness of registry data needed to calculate standard incidence statistics; the certification criteria do not address the accuracy or completeness of cancer treatment information (https://www.naaccr.org/certification-criteria/).

Women included in these analyses were diagnosed with a first primary Hodgkin lymphoma, non-Hodgkin lymphoma, breast cancer, thyroid cancer, ovarian cancer or cervical/uterine cancers at ages 15-39 during 2003-2014. Among 9,247 total cases identified in the CCR, we excluded those without Medicaid or private insurance at diagnosis (N=6,179). We further excluded those who were not continuously enrolled in insurance for at least one year following their cancer diagnosis (N=726), leaving 2,342 cases for analysis. Patient and tumor characteristics, including cancer type, age at diagnosis, date of diagnosis, SEER summary stage, (local, regional, distant, unstaged, unknown), and estrogen receptor status (breast cancer only: positive, negative, borderline) were determined using CCR data. Cancer types were categorized using an AYA-specific recode of International Classification of Diseases for Oncology 3rd Edition (ICD-O-3) primary site and histology codes.[6] Ovarian cancer was defined as germ cell and trophoblastic neoplasms of the gonads and carcinomas of the gonads.

Treatment information

We considered insurance claims data to be the gold standard on the assumption that treatment information in the claims is likely to be highly complete and accurate, given the low likelihood that insured patients would pay for cancer treatments out of pocket, or that medical facilities would charge insurers for services not performed. Medicaid and private insurance claims data were reviewed for the 12 months following cancer diagnosis. Codes from the International Classification of Diseases, Ninth Revision (ICD-9), Healthcare Common Procedure Coding System (HCPCS), Common Procedural Terminology (CPT), National Drug Codes (NDCs), and hospital revenue codes were used to identify cancer treatments in the claims data. The codes used to define treatment types (chemotherapy, radiation, hormone therapy) (Supplemental Table 1) were generated using lists from the Cancer Research Network published on the National Cancer Institute's website,[7] the report comparing SEER treatment information to Medicare claims,[1] clinical expertise, and collaborator input. Individuals were considered to have received chemotherapy, radiation, or

hormone therapy according to the claims data if at least one claim within the 12-months after diagnosis included a code for that treatment type.

The CCR collects information on first course of treatment, including chemotherapy, radiation, and hormone therapy, as well as dates of initiation of each of these therapies. The first course of treatment includes all methods recorded in the treatment plan and administered to the patient prior to disease progression or recurrence.[8] There is no defined interval for the collection of this information; first course of treatment is updated as it becomes available. For our analyses, CCR treatment information for each treatment type was categorized as received, not received, or missing/unknown.

Calculation

Using Medicaid and private insurance claims data as the gold standard, we calculated the sensitivity and positive predictive value (PPV) of chemotherapy, radiation, and hormone therapy information from the CCR. The sensitivity was used as a measure of completeness, and the PPV was used as a measure of accuracy. We also calculated the kappa coefficient and the percent agreement to assess concordance between the claims and CCR data. Kappa values were interpreted as follows: 0.0-0.20 as no agreement, 0.21-0.39 as minimal agreement, 0.40-0.59 as weak agreement, 0.60-0.79 as moderate agreement, 0.80-0.90 as strong agreement, and >0.90 as almost perfect agreement.[9] Because there is no defined interval for collection of treatment information by the registry, we considered both a 12-month and a 6-month postdiagnosis claims window (based on date of service in the claims) for comparison with CCR treatment data. Those with missing or unknown information for chemotherapy (2%), radiation (3%), or hormone therapy (5%) in the CCR were excluded from analyses for that treatment type. Comparisons of hormone therapy data were restricted to women diagnosed with breast cancer. Women with thyroid cancer were excluded from analyses of chemotherapy receipt, and those with ovarian cancer were excluded from analyses of radiation receipt, as few women with these cancer types were expected to have received these therapies. Stratified analyses were conducted according to cancer type, age at diagnosis, summary stage, year of diagnosis, and insurance type within the first 12 months after diagnosis.

To assess whether dates of treatment initiation were similar between the two data sources, we calculated the difference in days between the date recorded in the CCR and the date of service from the first claim. Those with dates that differed by more than 1 year between the two sources (N=2 for chemotherapy, N=2 for hormone therapy) were excluded from analyses comparing dates of treatment initiation.

Results

A total of 2,342 AYA women with cancer met the inclusion criteria and were included in these analyses. The most common cancer types were breast cancer (40%) and thyroid cancer (28%) (Table 1). The majority of women were diagnosed between the ages of 30 and 39 years (74%) with localized stage disease (52%), and had private insurance only (66%).

Using a 12-month postdiagnosis claims window, the sensitivity and PPV of the CCR data for chemotherapy were 86.4% and 81.9%, respectively, for all cancer types combined (Table 2). The kappa statistic (48.8%) and percent agreement (78.0%) indicated weak to moderate overall concordance between chemotherapy recorded in the claims data and in the CCR data. Across cancer types, the sensitivity ranged from 51.5% among women with cervical/uterine cancers to 96.5% among those with Hodgkin lymphoma, while the PPV ranged from 77.8% among women with cervical/uterine cancers to 82.6% among those with breast cancer. Little variation in sensitivity and PPV was observed according to age at diagnosis or type of insurance, but both measures were lower among women with localized disease than among those with regional or distant stage disease. No clear patterns were observed according to year of diagnosis, though sensitivity was >80% for all diagnosis year categories. Values of sensitivity tended to be slightly higher, and values of PPV were generally lower, when a 6-month postdiagnosis claims window was used (Supplemental Table 2).

The sensitivity of the CCR data for radiation receipt was 74.4% for all cancer types combined using a 12-month claims window, and ranged from 61.9% among women with Hodgkin lymphoma to 95.2% among those with cervical/uterine cancers (Table 3). Kappa and percent agreement were also highest for cervical/uterine cancers. The PPV was 83.0% for all cancer types combined and ranged from 78.4% for thyroid cancer to 89.7% for Hodgkin lymphoma. None of the measures examined varied greatly according to age at diagnosis, and no consistent patterns were observed according to year of diagnosis, though sensitivity was highest (78.8%) and PPV was lowest (70.0%) in the most recent years (2012-2014). In analyses according to insurance type, the sensitivity was slightly higher and the PPV was slightly lower, among women with private insurance only during the 12 months after diagnosis as compared to those with any Medicaid. Sensitivity, kappa, and percent agreement were all highest among women with localized stage disease, while PPV varied little by stage. When a 6-month postdiagnosis claims window was used, sensitivity was consistently higher and PPV was consistently lower compared to analyses using a 12-month window (Supplemental Table 3).

For women with breast cancer, the sensitivity and PPV for hormone therapy data from the CCR compared to the claims were 67.0% and 70.1%, respectively, when a 12-month claims window was used (Table 4). Concordance measures suggested minimal to weak agreement between the CCR data and the claims data. While the sensitivity was higher when a 6-month claims window was used, values of PPV, kappa, and percent agreement were substantially lower.

Among women with chemotherapy receipt recorded in both the CCR and the claims, the date of initiation in the CCR exactly matched the date of the first chemotherapy claim for 37% (Figure 1). Dates differed between the two sources by more than 30 days for 37%, with the majority of these having a date of service from their first claim that was later than the date of initiation in the registry. For radiation, 28% of women had dates in the CCR that exactly matched those of the first identified claim, while 61% had a date for the first claim that was 30 days or fewer before the date in the CCR. Only 7% of dates for the first radiation claim and radiation in the CCR differed by 30 or more days. Among women with breast cancer who had hormone therapy receipt in both the CCR and the claims data, 31% had

dates that matched exactly between the two sources, and 23% had dates that differed by 30 or more days.

Discussion

In this cohort of AYA women diagnosed with cancer in North Carolina, we assessed the validity of state cancer registry treatment data using insurance claims information as the gold standard. We found that chemotherapy information recorded in the registry was reasonably complete and accurate, with sensitivity and PPV both within the range of 80-90%. Though the PPV of registry data for radiation was similar to chemotherapy, the sensitivity was lower (74%), indicating some underascertainment of radiation receipt by the registry. Our analyses demonstrated low sensitivity and PPV for hormone therapy, suggesting that registry data alone could not be used to accurately estimate the proportion of breast cancer patients treated with endocrine agents during this time period.

Results of the current study for chemotherapy differ from those previously reported for data from SEER registries compared to Medicare claims. Among patients diagnosed with cancers of the bladder, female breast, colon or rectum, lung, ovary, pancreas or prostate at ages 65 or older, the sensitivity and PPV for chemotherapy data in SEER were 68% and 90%, respectively, using a 12-month postdiagnosis window for identification of chemotherapy in Medicare claims.[1] In contrast, among AYA women in North Carolina in the current study, the sensitivity was considerably higher (86.4%) and exceeded the PPV (81.9%).

Although we considered the claims data to represent the gold standard source of treatment information, it should be noted that some underascertainment of chemotherapy receipt in the claims data is possible. The list of drug and procedure codes that we used to identify chemotherapy in the claims incorporated codes from NCI lists,[7] the report comparing SEER treatment information to Medicare claims, and other sources, but there may be additional codes for chemotherapy that we did not include. Patients receiving care outside of their insurance plan, though likely an uncommon occurrence, could also lead to underascertainment in the claims. Given the differing patterns between our findings and those reported in SEER-Medicare,[1] additional investigation of chemotherapy accuracy and completeness in population-based registries is warranted.

Using a 12-month postdiagnosis claims window, we found moderately high values of PPV and sensitivity for radiation data recorded in the registry. However, these findings differed substantially from those observed using a 6-month claims window, in which the sensitivity was somewhat improved while the PPV was noticeably decreased. Differences between the 12-month and 6-month results appeared to be greatest for breast cancer and lymphoma, cancer types for which most patients in this age group also receive chemotherapy. Radiation typically occurs after chemotherapy in these patients, and may therefore begin too late to be accurately recorded by the registry. The greater underascertainment for radiation compared to chemotherapy may also be partially explained by a greater likelihood for radiation oncology centers to be freestanding and privately owned compared to chemotherapy clinics, which are more often part of the hospital; information on treatments received outside of the hospital setting may be more challenging to collect. Overall, these results suggest that

caution is warranted in studies using registry data alone to classify patients as treated or untreated with radiation.

Several prior studies have evaluated the accuracy and completeness of registry radiation data specifically among breast cancer patients, with highly variable results across registries.[2, 3, 10–13] Using Medicare claims as a the gold standard, a study of breast cancer patients aged 66 and older diagnosed in 2001-2007 reported sensitivities of radiation data ranging from 72.6% to 94.4% across SEER registries.[3] The same study also evaluated radiation data from three non-SEER registries compared to Medicare claims, finding sensitivities of 48.4%, 56.1%, and 81.1% for Florida, Texas, and New York, respectively. Our results, among a contemporary cohort of AYA women identified in a non-SEER registry, are within the range of those reported in other cohorts, and illustrate that underascertainment of radiation by population-based registries remains a concern in the breast cancer context.

In our data, registry information on hormone therapy for breast cancer appeared to be less accurate and complete than information on chemotherapy or radiation, a finding which may be explained by later initiation of hormone therapy than other therapies. While we used insurance claims data as the gold standard, other studies using either medical records or self-reported treatment information have reported similarly low values of sensitivity for hormone therapy data from population-based registries. In a study examining the quality of breast cancer treatment data in the Illinois State Cancer Registry for patients diagnosed between 2005 and 2008, the sensitivity of hormone therapy data in the registry was 62% compared to self-reported information, and 48% compared to the medical record.[2] Likewise, a study of breast cancer patients aged 65 and older in the New Mexico Tumor Registry reported a sensitivity of 59.7% for hormone therapy data from the registry compared to medical chart information recorded within 6 months after diagnosis.[14] Taken together, these findings, along with those of the current study, suggest that use of registry data to identify breast cancer patients treated with hormone therapy could result in considerable misclassification. Other data sources, such as insurance claims, should be used to augment registry data in breast cancer studies requiring hormone therapy information.

To our knowledge, ours is the first study to evaluate the validity of treatment initiation dates recorded in the registry. For radiation and hormone therapy, we found fairly high concordance between the registry and the claims, with <25% of women having dates of initiation that differed by more than 30 days between the two data sources. Though the majority of women also differed by 30 or fewer days for chemotherapy, it is unclear why more than one-third had a date for their first chemotherapy claim that was more than 30 days after the date of chemotherapy initiation recorded in the registry. These results may be informative in the design of algorithms for capturing cancer recurrence, in studies evaluating time to cancer treatment as a measure of cancer care quality, or other research activities using registry-based information on date of treatment initiation.

This study has some potential limitations. For our analysis, we considered insurance claims to be the gold standard source for information on cancer therapy receipt. It is possible that claims data could overestimate the proportion of patients receiving a particular treatment type, if claims generated within the first year of diagnosis reflected treatment for early

recurrence, rather than initial treatment for the primary cancer diagnosis. Recurrence information is not available from the CCR in North Carolina or most other states. Additionally, our analyses only included young women with cancer who linked with the insurance claims data; findings may not be generalizable to all cases in the registry. Analyses were constrained by limited sample size, which precluded cancer type-specific analyses of validity measures according to disease stage or other characteristics.

In summary, population-based cancer registries are important resources for research on cancer care, but it is critical to consider data quality when interpreting findings from registry-based studies. Results of the current study suggest that registry data is a reasonably accurate and complete source of information on chemotherapy receipt among young women with cancer. However, radiation and hormone therapy information from the registry may be best supplemented with data from other sources when attempting to identify patients treated with these therapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

[1]. Noone AM, Lund JL, Mariotto A, Cronin K, McNeel T, Deapen D, Warren JL, Comparison of SEER Treatment Data With Medicare Claims, Medical care 54(9) (2016) e55–64. [PubMed: 24638121]

[2]. Silva A, Rauscher GH, Ferrans CE, Hoskins K, Rao R, Assessing the quality of race/ethnicity, tumor, and breast cancer treatment information in a non-SEER state registry, Journal of registry management 41(1) (2014) 24–30. [PubMed: 24893185]

[3]. Walker GV, Giordano SH, Williams M, Jiang J, Niu J, MacKinnon J, Anderson P, Wohler B, Sinclair AH, Boscoe FP, Schymura MJ, Buchholz TA, Smith BD, Muddy water? Variation in reporting receipt of breast cancer radiation therapy by population-based tumor registries, International journal of radiation oncology, biology, physics 86(4) (2013) 686–93.

[4]. German RR, Wike JM, Bauer KR, Fleming ST, Trentham-Dietz A, Namiak M, Almon L, Knight K, Perkins C, Quality of cancer registry data: findings from CDC-NPCR's Breast and Prostate Cancer Data Quality and Patterns of Care Study, Journal of registry management 38(2) (2011) 75–86. [PubMed: 22096878]

[5]. Meyer AM, Olshan AF, Green L, Meyer A, Wheeler SB, Basch E, Carpenter WR, Big data for population-based cancer research: the integrated cancer information and surveillance system, North Carolina medical journal 75(4) (2014) 265–9. [PubMed: 25046092]

[6]. Surveillance, Epidemiology, and End Results Program. AYA Site Recode/WHO 2008 Definition. Available from: https://seer.cancer.gov/ayarecode/aya-who2008.html.

[7]. Cancer Research Network. Cancer Therapy Look-up Tables. Available from: http://www.hcsrn.org/crn/en/RESEARCH/LookupTables/.

[8]. North Carolina Central Cancer Registry. 2016 Cancer Collection and Reporting Manual (CCARM).

[9]. McHugh ML, Interrater reliability: the kappa statistic, Biochemia medica 22(3) (2012) 276–82. [PubMed: 23092060]

[10]. Du X, Freeman JL, Goodwin JS, Information on radiation treatment in patients with breast cancer: the advantages of the linked medicare and SEER data. Surveillance, Epidemiology and End Results, Journal of clinical epidemiology 52(5) (1999) 463–70. [PubMed: 10360342]

[11]. Virnig BA, Warren JL, Cooper GS, Klabunde CN, Schussler N, Freeman J, Studying radiation therapy using SEER-Medicare-linked data, Medical care 40(8 Suppl) (2002) Iv-49-54.

[12]. Jagsi R, Abrahamse P, Hawley ST, Graff JJ, Hamilton AS, Katz SJ, Underascertainment of radiotherapy receipt in Surveillance, Epidemiology, and End Results registry data, Cancer 118(2) (2012) 333–41. [PubMed: 21717446]

[13]. Malin JL, Kahn KL, Adams J, Kwan L, Laouri M, Ganz PA, Validity of cancer registry data for measuring the quality of breast cancer care, Journal of the National Cancer Institute 94(11) (2002) 835–44. [PubMed: 12048271]

[14]. Du XL, Key CR, Dickie L, Darling R, Delclos GL, Waller K, Zhang D, Information on chemotherapy and hormone therapy from tumor registry had moderate agreement with chart reviews, Journal of clinical epidemiology 59(1) (2006) 53–60. [PubMed: 16360561]

Highlights

- Cancer registry data on chemotherapy was fairly complete compared to claims data.

- Radiation and hormone therapy information in the registry appeared less complete.

- Our findings support using registry-based chemotherapy information in research.
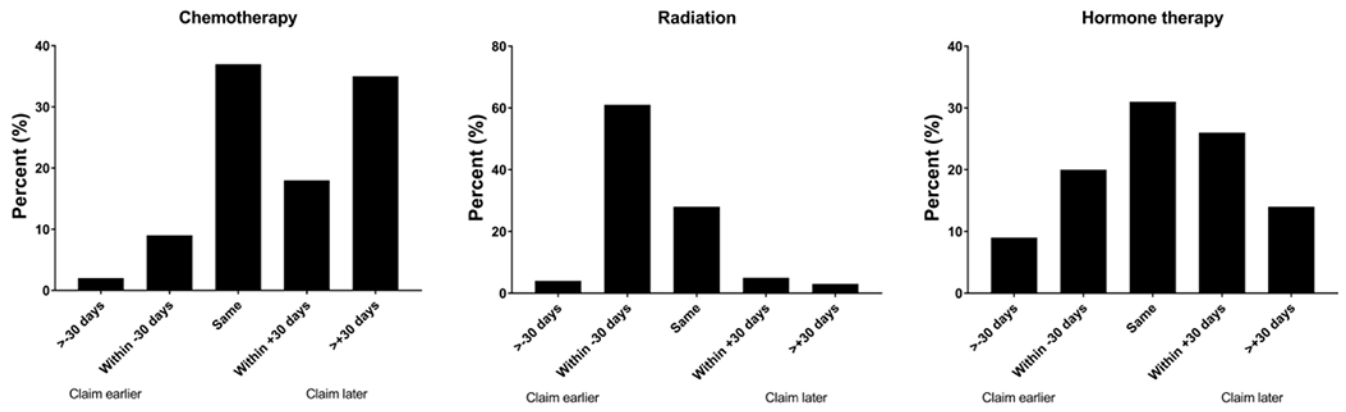
Figure 1.
Distribution of difference in days between date of treatment initiation in the claims data and in the registry data

Table 1.

Characteristics of adolescent and young adult women with cancer in North Carolina, 2003-2014

	N	%
Total	2,342	100
Cancer type		
Breast	942	40%
Hodgkin lymphoma	176	8%
Non-Hodgkin lymphoma	145	6%
Thyroid	651	28%
Cervical/uterus	327	14%
Ovarian	101	4%
Age at diagnosis		
15-29	612	26%
30-34	645	28%
35-39	1,085	46%
Stage		
Localized	1,218	52%
Regional	854	36%
Distant	231	10%
Unstaged/unknown	39	2%
Year of diagnosis		
2003-2005	529	23%
2006-2008	683	29%
2009-2011	612	26%
2012-2014	518	22%
Insurance type within 12 months post-diagnosis		
Any Medicaid	800	34%
Private insurance only	1,542	66%

Table 2.

Comparison of chemotherapy information in the registry and the claims using a 12-month claims window [a]

	Claims No, CCR No	Claims Yes, CCR No	Claims No, CCR Yes	Claims Yes, CCR Yes	Sensitivity	PPV	Kappa	% agreement
All	331	151	212	959	86.4 (84.4, 88.4)	81.9 (79.7, 84.1)	48.8 (44.2, 53.3)	78.0%
Cancer type								
Breast	107	61	131	622	91.1 (88.9, 93.2)	82.6 (79.9, 85.3)	39.8 (33.0, 46.7)	79.2%
Cervix/uterus	168	63	19	67	51.5 (43.0, 60.1)	77.9 (69.1, 86.7)	43.6 (33.8, 53.5)	74.1%
Hodgkin lymphoma	1	5	30	137	96.5 (93.5, 99.5)	82.0 (76.2, 87.9)	0 (−10.5, 9.7)	79.8%
Non-Hodgkin lymphoma	25	19	18	81	81.0 (73.3, 88.7)	81.8 (74.2, 89.4)	38.9 (22.6, 55.2)	74.1%
Ovarian	30	3	14	52	94.6 (88.5, 100.0)	78.8 (68.9, 88.7)	64.3 (49.4, 79.3)	82.8%
Age at diagnosis								
15-29	61	31	51	220	87.7 (83.6, 91.7)	81.2 (76.5, 85.8)	44.3 (34.2, 54.4)	77.4%
30-34	91	38	51	272	87.7 (84.1, 91.4)	84.2 (80.2, 88.2)	53.1 (44.6, 61.7)	80.3%
35-39	179	82	110	467	85.1 (82.1, 88.1)	80.9 (77.7, 84.1)	48.1 (41.8, 54.4)	77.1%
Stage								
Localized	268	115	76	292	71.7 (67.4, 76.1)	79.4 (75.2, 83.5)	49.2 (43.0, 55.4)	74.6%
Regional	29	17	102	507	96.8 (95.2, 98.3)	83.3 (80.3, 86.2)	25.0 (16.0, 33.9)	81.8%
Distant	16	11	31	159	93.5 (89.8, 97.2)	83.7 (78.4, 88.9)	32.6 (17.2, 48.0)	80.6%
Year of diagnosis								
2003-2004	78	32	26	271	89.4 (86.0, 92.9)	91.3 (88.0, 94.5)	63.2 (54.6, 71.8)	85.7%
2005-2007	105	29	90	251	89.6 (86.1, 93.2)	73.6 (68.9, 78.3)	45.7 (37.6, 53.7)	74.9%
2008-2010	87	55	38	257	82.4 (78.1, 86.6)	87.1 (83.3, 90.9)	49.9 (41.2, 58.7)	78.7%
2011-2013	61	35	58	180	83.7 (78.8, 88.7)	75.6 (70.2, 81.1)	36.6 (26.1, 47.0)	72.2%
Insurance type within 12 months post-diagnosis								
Any Medicaid	115	54	81	375	87.4 (84.3, 90.6)	82.2 (78.7, 85.7)	47.9 (40.4, 55.4)	78.4%
Private insurance only	216	97	131	584	85.8 (83.1, 88.4)	81.7 (78.8, 84.5)	49.2 (43.5, 54.9)	77.8%

[a] Excludes those with missing chemotherapy information from the registry (N=38)

Table 3.

Comparison of radiation information in the registry and the claims using a 12-month claims window[a]

	Claims No, CCR No	Claims Yes, CCR No	Claims No, CCR Yes	Claims Yes, CCR Yes	Sensitivity	PPV	Kappa	% agreement
All	927	277	165	803	74.4 (71.8, 77.0)	83.0 (80.6, 85.3)	59.3 (55.9, 62.7)	79.7%
Cancer type								
Breast	314	150	66	369	71.1 (67.2, 75.0)	84.8 (81.5, 88.2)	52.2 (46.7, 57.7)	76.0%
Cervix/uterus	227	4	12	79	95.2 (90.6, 99.8)	86.8 (80.0, 93.8)	87.4 (81.4, 93.4)	95.0%
Hodgkin lymphoma	80	32	6	52	61.9 (51.5, 72.3)	89.7 (81.8, 97.5)	55.1 (43.2, 67.1)	77.6%
Non-Hodgkin lymphoma	94	16	5	28	63.6 (49.4, 77.9)	84.9 (72.6, 97.1)	63.0 (48.8, 77.1)	85.3%
Thyroid	212	75	76	275	78.6 (74.3, 82.9)	78.4 (74.0, 82.7)	52.2 (45.6, 58.9)	76.3%
Age at diagnosis								
15-29	233	83	43	202	70.9 (65.6, 76.2)	82.5 (77.7, 87.2)	55.2 (48.4, 62.0)	77.5%
30-34	270	63	54	209	76.8 (71.8, 81.9)	79.5 (74.6, 84.4)	60.3 (53.9, 66.8)	80.4%
35-39	424	131	68	392	75.0 (71.2, 78.7)	85.2 (82.0, 88.5)	60.9 (56.1, 65.7)	80.4%
Stage								
Localized	653	97	65	334	77.5 (73.6, 81.4)	83.7 (80.1, 87.3)	69.5 (65.1, 73.8)	85.9%
Regional	143	150	88	414	73.4 (69.8, 77.1)	82.5 (79.1, 85.8)	32.7 (25.9, 39.5)	70.1%
Distant	104	26	10	52	66.7 (56.2, 77.1)	83.9 (74.7, 93.0)	59.8 (48.3, 71.4)	81.3%
Year of diagnosis								
2003-2004	188	75	23	202	72.9 (67.7, 78.2)	89.8 (85.8, 93.7)	60.3 (53.4, 67.1)	79.9%
2005-2007	272	102	33	231	69.4 (64.4, 74.3)	87.5 (83.5, 91.5)	58.0 (51.9, 64.2)	78.8%
2008-2010	254	58	42	214	78.7 (73.8, 83.5)	83.6 (79.1, 88.1)	64.6 (58.4, 70.9)	82.4%
2011-2013	213	42	67	156	78.8 (73.1, 84.5)	70.0 (63.9, 76.0)	53.9 (46.3, 61.4)	77.2%
Insurance type within 12 months post-diagnosis								
Any Medicaid	311	113	46	275	70.9 (66.4, 75.4)	85.7 (81.8, 89.5)	57.6 (51.8, 63.3)	78.7%
Private insurance only	616	164	119	528	76.3 (73.1, 79.5)	81.6 (78.6, 84.6)	60.2 (56.1, 64.4)	80.2%

[a]Excludes those with missing radiation information from the registry (N=69)

Table 4.

Comparison of hormone therapy information in the registry and the claims

		Claims No, CCR No	Claims Yes, CCR No	Claims No, CCR Yes	Claims Yes, CCR Yes	Sensitivity	PPV	Kappa	% agreement
All breast [a]									
12 mo		331	105	91	213	67.0 (61.8, 72.2)	70.1 (64.9, 75.2)	45.7 (39.2, 52.2)	74%
6 mo		404	32	218	86	72.9 (64.9, 80.9)	28.3 (23.2, 33.4)	23.1 (17.0, 29.2)	66%
ER positive breast [b]									
12 mo		68	94	82	204	68.5 (63.2, 73.7)	71.3 (66.1, 76.6)	13.5 (4.2, 22.9)	61%
6 mo		133	29	201	85	74.6 (66.6, 82.6)	29.7 (24.4, 35.0)	9.6 (3.1, 16.2)	49%

[a]Excludes those with missing ER status (N=142) or missing hormone therapy information from the registry (N=60)

[b]Excludes those missing hormone information from the registry (N=49)