DATA ARCHITECTURE AND MODELS

Consistency Between State's Cancer Registry and All-Payer Claims Database in Documented Radiation Therapy Among Patients Who Received Breast Conservative Surgery

Chenghui Li, PhD[1]; Sindhu J. Malapati, MD[2]; John T. Guire, CTR[3]; and Laura F. Hutchins, MD[2]

PURPOSE Arkansas is one of only four known states that have linked All-Payer Claims Database (APCD) to state's cancer registry (Arkansas Cancer Registry [ACR]). We evaluated the reporting consistency of radiation therapy (RT) between the two sources.

METHODS Women age ≥ 18 years diagnosed in 2013-2017 with early-stage hormone receptor–positive breast cancer who received breast-conserving surgery were identified. Patients must have continuous insurance coverage (any private plans, Medicaid, and Medicare) in the 13 months (month of diagnosis and 12 months after). Receipt of RT was identified independently from ACR and APCD. We calculated sensitivity, specificity, positive predictive value, and negative predictive value for receipt of RT coded by the registry compared with APCD billing claims as the gold standard. We assessed the degree of concordance between the data sources by Cohen's kappa statistics.

RESULTS The final sample included 2,695 patients who were in both databases and satisfied our inclusion/exclusion criteria. Using APCD as the gold standard, there were high sensitivity (88.1%) and positive predictive value (87.7%) and moderate specificity (71.1%) and negative predictive value (71.8%). The overall agreement between the two sources was 83.0%, with a kappa statistic of 0.59 (95% CI, 0.56 to 0.63). Consistency measures varied by age, stage, and insurance type with Medicare fee-for-service coverage only having the best and private insurance only the worse consistency.

CONCLUSION In patients with early-stage hormone receptor–positive breast cancer who received breast-conserving surgery, recording of RT receipt was moderately consistent between Arkansas APCD and ACR. Future studies are needed to identify factors affecting reporting consistency to better use this unique resource in addressing population health problems.

INTRODUCTION

All-payer claims databases (APCDs) are large state databases that include medical, pharmacy, and dental claims and eligibility and provider files collected from private and/or public payers.[1] As an effort to increase price transparency, currently, 18 states have legislation mandating the creation and use of APCDs or are actively establishing APCDs and more than 30 states are maintaining or developing an APCD or have a strong interest in developing an APCD.[2] State APCDs may include individuals of all ages and can capture longitudinally health care utilization across settings and payers, providing a more complete picture of care even when patients switch insurance plans or are covered concurrently by multiple plans or payers.[2] Arkansas is a pioneer of state APCD. Arkansas' APCD included claims from commercial plans, State Employee plans, Medicaid, and Medicare, covering 80% of Arkansans. Although many states have

developed APCD, Arkansas is one of only a few states that have linked APCD to state's central cancer registry,[3] which provides the critical tumor-related information for cancer research.

Although APCD offers many potential advantages over one-payer claims databases, many challenges have been identified, which could threaten the completeness of data on treatment patterns.[4] Therefore, it is important to assess the quality of the linked database. Only one other state has compared reported cancer treatments in the linked APCD registry database. Utah has linked their APCD to state's registry and found that APCD improves cancer treatment data collection; however, identifying treatment using both databases increased false-positive rate when compared with manual abstraction as the gold standard.[5]

In this study, we assessed the quality of reporting radiation therapy (RT) in the linked Arkansas APCD-Arkansas

CONTEXT

Key Objective

To evaluate reporting consistency of radiation therapy (RT) between Arkansas' All-Payer Claims Database (APCD) and Arkansas Cancer Registry in adult women with early-stage hormone receptor–positive breast cancer who received breast-conserving surgery.

Knowledge Generated

Recording of RT receipt was moderately consistent between APCD and Arkansas Cancer Registry. Consistency measures varied by age, stage, and insurance type with Medicare. Fee-for-service coverage only had the best consistency, and private insurance only the worse consistency.

Relevance

Cancer registries only documented first course of treatment and may under-report nonsurgery treatment. APCD can track patients longitudinally across settings and payers throughout cancer care continuum and may improve documentation of treatment among patients with continuous insurance coverage. Our findings showed good consistency between the two sources, and APCD can be a valuable resource to determine RT received as part of first-course treatment by patients with early-stage hormone receptor–positive breast cancer who received breast-conserving surgery.

Cancer Registry (ACR) database among females with early-stage hormone receptor–positive (HR+) breast cancer (BC) who received breast-conserving surgery (BCS). Multiple randomized trials comparing BCS with mastectomy showed no differences in overall survival.[6-10] Radiation after BCS is recommended[11] and has been shown to significantly reduce the rates of ipsilateral breast tumor recurrence and improve BC-specific and overall survival.[12] The specific aims of the study were to (1) compare the consistency in documented RT receipt between APCD and ACR and (2) identify factors associated with discordance between the two sources.

METHODS

This study used only deidentified secondary databases and was determined by our institution's institutional review board not to be a human subject study (IRB# 274492).

Data Sources

We used Arkansas APCD linked to ACR, established by the Arkansas Center for Health Improvement, a nonpartisan, independent health policy center dedicated to improving the health of Arkansans.[13] Enrollee's last name, birth date, and sex were used to achieve this linkage, with an estimated matching rate of 96.5%.[14] This study included BC patients diagnosed from 2013 to 2017 in ACR and claims from January 1, 2013, to December 31, 2019.

Patient Selection

From ACR, we identified Arkansas women age ≥ 18 years diagnosed in 2013-2017 with early-stage (stages 0-3) HR+ (ie, estrogen receptor– and/or progesterone receptor–positive) BC (ICD-O: C50.0x-C50.9x) who received BCS. Stage was based on either pathologic or clinical stage, whichever was higher. BCS recorded in ACR was defined as partial or less than total mastectomy with/without dissection of axillary lymph nodes. Patients were excluded if they (1) received no surgery, received mastectomy, or had

unknown surgery status; (2) were HR– or had unknown HR status; (3) had multiple cancer diagnoses in ACR; (4) were diagnosed from autopsy or death certificate; or (5) had stage 4 BC or unknown stage. Date of cancer diagnosis was from ACR, excluding patients with missing year and/or month of diagnosis. For those missing only day of diagnosis, we used the first day of the recorded diagnosis month as the day of diagnosis.

We then used the unique alias patient IDs created by the Arkansas Center for Health Improvement to identify patients' claims in APCD. The following additional inclusion criteria were implemented to ensure complete claims for all individuals in the study. For private plans and Medicaid, we required continuous medical insurance coverage in the month of BC diagnosis and 12 months after, for a total of 13 months of continuous coverage. Medicare claims were from the Centers for Medicare & Medicaid Services and do not include claims from Medicare Advantage (MA) plans. Therefore, for Medicare claims, we required continuous part A and B coverage and no MA coverage (ie, hereafter Medicare fee-for-service [FFS]) throughout the 13 months. However, claims from MA plans can be part of Arkansas APCD private plan claims if these plans were administered through insurers in Arkansas. For this study, we consider them as private plans.

Receipt of RT

Receipt of RT was identified independently from ACR and APCD to facilitate comparison. ACR collects information on the first course of therapy as documented in the treatment plan. Six variables in ACR could inform RT receipt: RADDOSE (regional dose [cGy]), RADMODALITY (regional RX modality), RADVOL (treatment volume/anatomic site), REASONNORAD (reason for no radiation), RXDATERAD (first date of radiation), and RXDTRADEND (end date of radiation), and the information can be slightly different across variables. We

consider a patient as having received RT if any of the six variables reported it or not if none of these variables reported it including no radiation or unknown radiation status.

In APCD, RT was identified using procedure codes for radiation delivery (Appendix Table A1). After previous studies,[5,15,16] we chose a 1-year postdiagnosis window to identify RT as part of the first course of treatment in APCD. We considered a patient as having received RT if she had at least one claim with these procedure codes. For sensitivity analyses, we used windows of 6 months and 2 years to look for RT claims. In addition, we relaxed the continuous insurance coverage requirement.

Radiation to Breast

In ACR, the variable RADVOL identifies the volume or anatomic target of the most clinically significant regional radiation delivered to the patient during the first course of therapy.[17] We consider a patient who received radiation to breast if the site was 18 breast (the primary target is the intact breast, and no attempt has been made to irradiate the regional lymph nodes) or 19 breast/lymph nodes (a deliberate attempt has been made to include regional lymph nodes in the treatment of an intact breast).[17]

In APCD, radiation procedure codes are not specific to any anatomic site. To identify radiation to breast, we required a diagnosis of BC (malignant neoplasm of female breast: ICD9 174.xx/ICD10 C50.xx; carcinoma in situ of breast: ICD9 233.0x/ICD10 D05.xx) on the day of the procedure to minimize the probability of identifying radiation directed to a metastatic site or for other cancers.

Radiation Starting Date

ACR records the first date when RT was delivered as part of the first course of therapy (RXDATERAD). In APCD, we used the date of the first radiation delivery claim within 1 year after BC diagnosis as the RT starting date.

Key Covariates

Covariates included patient's age at diagnosis ($<$ 64, 64-65, and $>$ 65 years), race/ethnicity (Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Others), stage (0-3), insurance type during the 13 months (private plans only, Medicaid only, Medicare FFS only, private/Medicaid, private/Medicare FFS, Medicaid/Medicare FFS, or all three), residence at time of diagnosis (rural/urban), and year of diagnosis. Age group 64-65 years was studied separately because insurance coverage during this time is expected to change because of age-related Medicare eligibility. Rurality was based on the 2013 National Center for Health Statistics' continuum codes of rurality.[18]

Analysis

We calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for RT receipt documented in the registry compared with APCD billing claims as the gold standard. We chose APCD as the gold standard because we required all patients to have continuous insurance coverage during the 13 months; thus, APCD should capture all RT received during this time. Billing claims were also used as the gold standard in previous studies evaluating quality of RT information in cancer registries.[15,16] Nonetheless, PPV and NPV when using APCD as the gold standard are the same as sensitivity and specificity when ACR is used as the gold standard.

In addition, we calculated the proportion of patients having records in agreement (ACR yes and APCD yes; ACR no and APCD no) and assessed the degree of agreement using Cohen's kappa statistic.[20] Both statistics do not assume a gold standard. The following scale was used to interpret kappa values: no agreement: $\leq$ 0, slight agreement: 0.01-0.20, fair agreement: 0.21-0.40, moderate agreement: 0.41-0.60, substantial agreement: 0.61-0.80, and almost perfect agreement: 0.81-1.00.[21] We also conducted these analyses for radiation to breast.

We assessed the concordance in the recorded first day of RT by calculating the proportion of patients with reported RT initiation (1) on the same day, (2) 1-7 days apart, (3) 8-15 days apart, (4) 16-30 days apart, (5) 31-60 days apart, (6) 61-90 days apart, or (7) 91-180 days apart among patients with documented RT in both sources. We also determined the sensitivity, specificity, PPV, NPV, proportion in agreement, and kappa statistics separately by the key covariates and performed a logistic regression of reporting disagreement (ref = in agreement) on the key covariates.

All analyses were conducted using SAS software, version 9.4 (SAS Institute, Cary, NC) and Stata software, version 17 (StataCorp LLC, College Station, TX). $P <$ .05 for a two-sided test is considered statistically significant.

RESULTS

The final cohort included 2,695 patients. Figure 1 shows patients excluded because of various selection criteria. Half of the patients were diagnosed at age $\geq$ 66 years, and 7% were diagnosed around the Medicare eligibility age (64-65 years). The majority of patients were White (86%) and living in rural areas (53%). During the 13 months, 67% had one type of coverage (Medicare FFS only: 32%, Medicaid only: 3%, and private plans only: 32%) and 33% had a combination of coverages.19%, 53%, 25%, and 3% of the patients had stage 0-3 diseases, respectively (Table 1).

Using claims within 1 year after diagnosis, we found 83.0% agreement in documented RT receipt (ACR yes, APCD yes, 61.7%; ACR no and APCD no, 21.3%). Using APCD as the gold standard, sensitivity, specificity, PPV, and NPV were 88.1%, 71.1%, 87.7%, and 71.8%, respectively. The kappa statistic was 0.59 (95% CI, 0.56 to 0.63). Using claims in 2 years did not significantly affect these statistics. However, using claims in 6 months significantly reduced specificity (54.6%), overall agreement (74.2%), and kappa statistic (0.45; 95% CI, 0.42 to 0.48). Not requiring 13 months of continuous coverage further reduced

specificity (48.7%), overall agreement (70.7%), and kappa statistic (0.38; 95% CI, 0.35 to 0.41; Table 2).

Consistency measures varied by age at diagnosis, with patients diagnosed at age ≥ 66 years having the highest specificity (90.7%) and PPV (95.1%), whereas specificity (44.4%) and PPV (80.2%) were lowest in 18-63 years age group. Across insurance groups, patients having Medicare FFS–only coverage in the 13 months had the best consistency with the highest sensitivity (89.0%), specificity (93.8%), PPV (97.1%), second best NPV (78.3%), and highest overall agreement (90.4%) and kappa statistic (0.78; 95% CI, 0.74 to 0.83); patients with private insurance only had the lowest specificity (45.7%), PPV (75.7%), overall agreement (73.8%), and kappa statistic (0.37; 95% CI, 0.31 to 0.73). By stage, patients with stage 0 cancer had the most consistency. Consistency differed by race/ethnicity, with

non-Hispanic White patients having the most consistent records; however, estimates for Hispanic patients and non-Hispanic Others may be unreliable because of small numbers (Table 2).

Of all RT claims (n = 106,989) within 1 year after diagnosis, nearly all (98.6%) claims had a BC diagnosis (Medicaid: 97.4%, private: 98.6%, and Medicare: 98.9%) as either the principal (Medicaid: 70.1%, private: 68.2%, and Medicare: 63.7%) or a secondary diagnosis. Among claims with BC as a secondary diagnosis, encounter for antineoplastic RT codes (ICD9 V58.0x, ICD10 Z51.0x) were most frequently used as the principal diagnosis (Medicaid: 94.2%, private: 97.5%, and Medicare: 96.7%).

Of the 1,896 patients with documented RT in the ACR, 92.4% were documented as being received radiation to

FIG 1. Patient selection flowchart. APCD, All-Payer Claims Database; ER, estrogen receptor; NOS, not otherwise specified; PR, progesterone receptor.
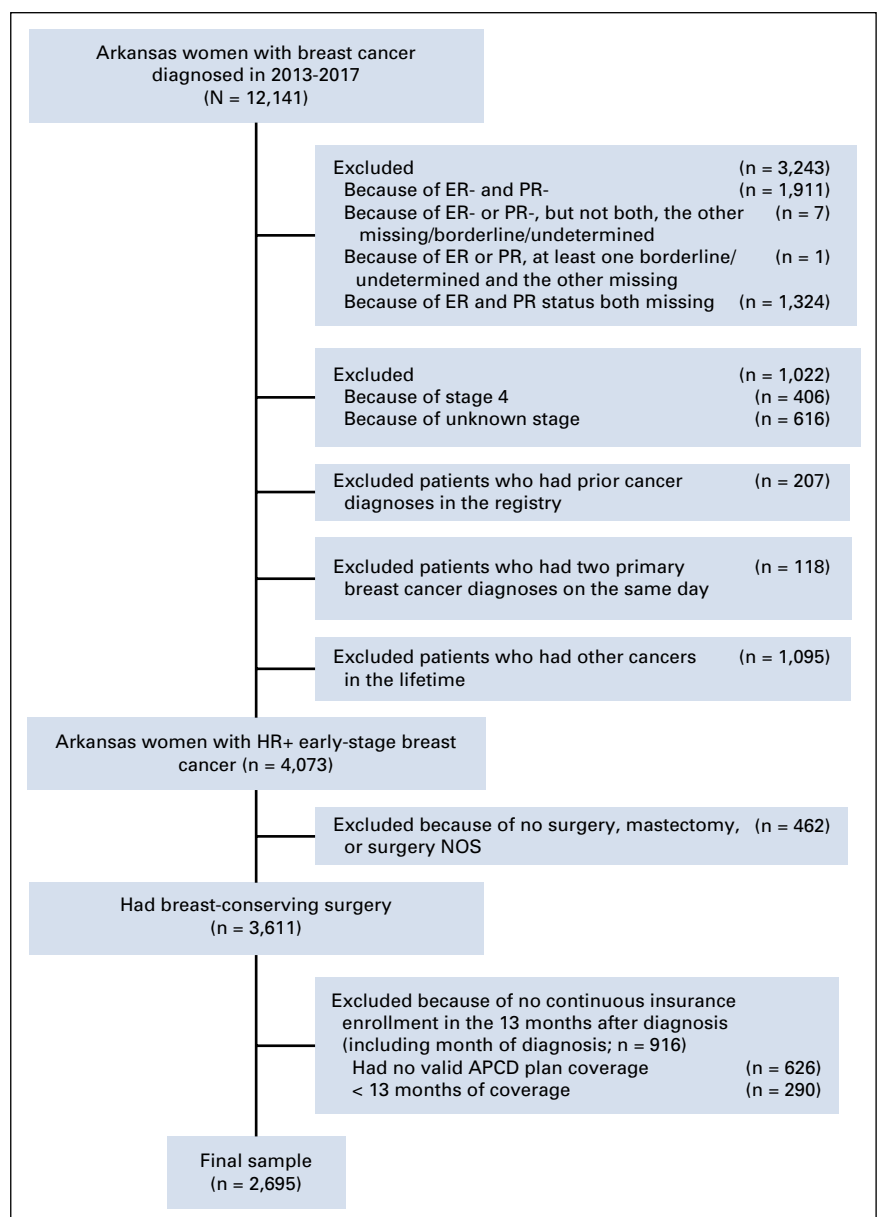
TABLE 1. Patient Characteristics

Characteristics	Total No. of Patients n = 2,695, No. (%)
Age group at diagnosis, years	
18-63	1,076 (40)
64-65	200 (7)
66 or above	1,419 (53)
Race category	
Hispanic	38 (1)
Non-Hispanic White	2,305 (86)
Non-Hispanic Black	311 (12)
Non-Hispanic Others	34 (1)
Unknown	7 (0.3)
Rural/urban	
Missing	4 (0.1)
Urban	1,258 (47)
Rural	1,433 (53)
Insurance in the 13 months	
Medicare FFS only	865 (32)
Medicaid only	79 (3)
Private only	874 (32)
Medicare and Medicaid	136 (5)
Private and Medicare	408 (15)
Private and Medicaid	273 (10)
All three	60 (2)
Stage (clinical or pathologic staging)	
0	502 (19)
1	1,421 (53)
2	685 (25)
3	87 (3)
Breast-conserving surgery	
Partial mastectomy, NOS	231 (9)
Partial mastectomy, without nipple resection	18 (1)
Lumpectomy/excisional biopsy	1,996 (74)
Re-excision of the biopsy site for gross/microscopic residual disease	407 (15)
Segmental mastectomy (including wedge resection/quadrantectomy/tylectomy)	43 (2)
Year of diagnosis	
2013	459 (17)
2014	505 (19)
2015	551 (20)
2016	592 (22)
2017	588 (22)

Abbreviation: FFS, fee-for-service; NOS, not otherwise specified.

breast and/or regional lymph nodes. Among the 1,888 patients with RT claims within 1 year after cancer diagnosis, 99.8% had claims with BC as a diagnosis. The overall agreement was 80.0% (ACR breast yes and APCD breast yes, 57.4%; ACR breast no and APCD breast no, 22.5%). The sensitivity, specificity, PPV, and NPV were 82.2%, 74.8%, 88.4%, and 64.4% respectively, with a kappa statistic of 0.55 (95% CI, 0.52 to 0.58; Table 3).

The mean documented RT starting date was 113 days after BC diagnosis in the ACR (n = 1,749) and 111 days in the APCD (n = 1,888). Of the 1,548 patients with a nonmissing starting date in both databases, the mean difference was only 2 days between the two sources; 92.0% had starting dates within 30 days apart (same day, 34.0%; 1-7 days apart, 39.1%; 8-15 days apart, 12.4%; and 16-30 days apart, 6.5%; Fig 2).

Using logistic regression on disagreement between the two sources, younger age at diagnosis versus age ≥ 66 years (odds ratio; 95% CI: 18-63 years, 1.56; 1.15 to 2.11 and 64-65 years, 1.60; 1.07 to 2.40), stage 1 versus 0 cancer (1.38; 1.03 to 1.85), and private insurance only (2.45; 1.73 to 3.47) versus Medicare FFS only in the 13 months were independently associated with a higher risk of disagreement (Table 4).

DISCUSSION

To our knowledge, we conducted the first study assessing RT quality in Arkansas's APCD linked with state's cancer registry among females with early-stage HR+ BC who received BCS. Overall, we found good concordance between APCD and ACR, with high sensitivity and PPV and moderate specificity and NPV. Medicare FFS claims had the best consistency with the registry, whereas private plan claims, which included claims from multiple private payers, had the worst consistency. Only two other states have linked state's APCD and cancer registry to facilitate cancer-related research.[5,22] APCDs can complement state cancer registry beyond the first course of treatment to track patients longitudinally for subsequent care throughout cancer survivorship and assess patterns and quality of care and outcomes.

Our findings are generally consistent with previous findings. We conducted a review of published studies of registry-claims–linked databases that have compared RT documented in cancer registry with claims when BC was either the only cancer site[15,19,23,24] or one of several cancer sites studied[5,16,25,26] (Appendix Table A2). On the basis of reported data in these studies, we calculated sensitivity, specificity, PPV, and NPV using claims as the gold standard and proportion in agreement and kappa statistic (if not already reported in a study and feasible) to facilitate comparison with our findings. Of the eight registry-claims–linked studies identified, five assessed SEER-Medicare–linked database.[15,16,19,25,26] These studies varied by years of data used (therefore varying number of SEER sites included), the follow-up window for identifying RT claims, and codes used to identify RT claims. Overall, there was a good agreement between SEER registries and Medicare claims, with 84.3%-94.2% agreement and kappa statistics ranging from

TABLE 2. Reporting Consistency of Radiation Therapy Between ACR and APCD Using APCD as the Gold Standard

Analysis	No.	Sensitivity, %	Specificity, %	PPV, %	NPV, %	% in Agreement	Kappa	(95% CI)
Main analysis								
Claims within one year of diagnosis	2,695	88.1	71.1	87.7	71.8	83.0	0.59	(0.56 to 0.63)
Sensitivity analysis								
Claims within two years of diagnosis	2,695	87.6	71.1	87.9	70.5	82.7	0.59	(0.55 to 0.62)
Claims within 6 months of diagnosis	2,695	88.5	54.6	72.8	77.6	74.2	0.45	(0.42 to 0.48)
Not requiring continuous insurance coverage in 13 months	3,611	87.9	48.7	68.7	75.9	70.7	0.38	(0.35 to 0.41)
By patient characteristics								
Age at diagnosis, years								
18-63	1,076	88.5	44.4	80.2	60.3	76.0	0.36	(0.30 to 0.42)
64-65	200	90.0	47.5	87.3	54.3	81.5	0.39	(0.23 to 0.56)
66 or above	1,419	87.4	90.7	95.1	77.8	88.5	0.75	(0.71 to 0.79)
Rural/urban residence at diagnosis (excluded missing)								
Urban	1,258	88.5	66.6	85.4	72.3	81.6	0.56	(0.51 to 0.61)
Rural	1,433	87.7	75.8	89.9	71.5	84.3	0.62	(0.58 to 0.67)
Race/ethnicity (excluded unknown)								
Hispanic[a]	38	92.6	54.5	83.3	75.0	81.6	0.51	(0.20 to 0.82)
Non-Hispanic White	2,305	88.3	73.4	88.8	72.4	83.9	0.61	(0.58 to 0.65)
Non-Hispanic Black	311	86.4	61.9	83.3	67.4	78.8	0.49	(0.39 to 0.60)
Non-Hispanic Others[a]	34	84.2	53.3	69.6	72.7	70.6	0.39	(0.08 to 0.69)
Stage								
0	502	86.6	83.9	88.1	81.9	85.2	0.7	(0.64 to 0.77)
1	1,421	87.6	68.7	87.5	68.9	82.1	0.56	(0.52 to 0.61)
2	685	90.3	60.5	87.2	67.5	83.4	0.53	(0.45 to 0.60)
3[a]	87	85.5	77.8	93.7	58.3	83.7	0.56	(0.36 to 0.76)
Insurance type in the 13 months								
Medicare only	865	89.0	93.8	97.1	78.3	90.4	0.78	(0.74 to 0.83)
Medicaid only[a]	79	83.3	73.7	90.9	58.3	81.0	0.52	(0.31 to 0.73)
Private only	874	88.5	45.7	75.7	67.5	73.8	0.37	(0.31 to 0.44)
Medicare and Medicaid[a]	136	87.2	91.4	93.2	84.1	89.0	0.78	(0.67 to 0.88)
Medicare and private	408	87.6	81.2	93.4	68.3	86.0	0.65	(0.56 to 0.73)
Private and Medicaid	273	87.4	58.8	90.2	51.7	82.1	0.44	(0.31 to 0.57)
All three[a]	60	85.0	80.0	89.5	72.7	83.3	0.63	(0.43 to 0.84)
Year of diagnosis								
2013	459	88.0	66.1	88.2	65.5	82.4	0.54	(0.45 to 0.63)
2014	505	86.7	71.3	88.5	68.0	82.4	0.57	(0.49 to 0.65)
2015	551	90.0	70.9	87.2	76.3	84.0	0.62	(0.55 to 0.69)
2016	592	87.0	69.3	87.0	69.3	81.8	0.56	(0.50 to 0.64)
2017	588	88.7	75.8	87.8	77.3	84.4	0.65	(0.58 to 0.71)

Abbreviations: ACR, Arkansas Cancer Registry; APCD, All-Payer Claims Database; NPV, negative predictive value; PPV, positive predictive value.
[a]Estimates may not be reliable because of low cell counts (< 10).

0.54 to 0.87 across the studies, which are consistent with our findings of Medicare FFS claims compared with ACR. However, agreement varied by SEER site,[15,16,19] cancer stage,[15,16,19] and cancer site.[16,26] Moreover, since only Medicare claims were used, these studies were limited to patients age ≥ 65 years with Medicare FFS coverage.

TABLE 3. Consistency in Reporting Radiation to Breast

ACR Radiation to Breast	Radiation Delivery Claims in APCD With a Breast Cancer Diagnosis		
	Yes	No	Total
ACR radiation to breast			
Yes	1,548	204	1,752
No	336	607	943
Total	1,884	811	2,695

Abbreviations: ACR, Arkansas Cancer Registry; APCD, All-Payer Claims Database.

Only three studies have assessed consistency of cancer registries with claims other than Medicare and included patients younger than age 65 years.[5,23,24] All three studies used state-specific data. Two studies used earlier data and compared state's cancer registries with claims from a large private payer, Blue Cross/Blue Shield (BCBS) in their states, both restricted to patients under age 65 years.[23,24] Doebbeling et al[23] matched Iowa cancer registry with Iowa's BCBS for women with in situ or invasive BC from 1989 to 1996. Claims for incident cancer patients in 1994 (n = 445) were examined for consistency with cancer registry; an agreement of 87% was reported for RT. Hillner et al linked claims from Virginia's BCBS with state's cancer registry from 1989 to 1991 and included 918 female patients with BC without in situ or metastatic disease.[24] Using claims as the gold standard, the sensitivity, specificity, PPV, and NPV for RT were 89.3%, 93.3%, 82.8%, and 96.0% respectively; 92% were in agreement with a kappa statistic of 0.81 (95% CI, 0.76 to 0.85).[24] Similar to Iowa's study,[23] no fixed window after surgery was specified to search for RT, which may identify subsequent treatment that was not part of the first course of therapy.[24]

Only one study compared between cancer registry and APCD. The study used Utah's SEER linked with APCD, covering approximately 90% of commercially insured population in Utah.[5] Across all cancer sites, agreement for RT was 88.6%. Using APCD as the gold standard, sensitivity, specificity, PPV, and NPV were 85.9%, 89.3%, 70.1%, and 95.6%, respectively; APCD identified additional 4.4% of patients with RT who were documented as no RT in Utah's SEER.[5] For 198 patients with BC under age 65 years, manual abstraction by certified tumor registrar was made using SEER's Patterns of Care abstraction form.[5] Compared with manual extraction, SEER demonstrated excellent sensitivity (95.4%) and specificity (100%). Adding APCD improved sensitivity (97.7%) but reduced specificity (85.5%). Using APCD only had moderate sensitivity (74.6%) and high specificity (85.7%).[5]

We found that risk of disagreement in RT differed by age at diagnosis and was lowest among patients age ≥ 66 years. Previous studies also found that concordance varies by age but were limited to patients age ≥ 65 years.[15,16,19] Only the Utah study included all ages.[5] Although separate consistency measures by age were not reported, the likelihood of identifying additional patients with RT claims in APCD missed by the registry was lower in younger groups (< 50 and 50-64 years) compared with those age ≥ 75 years, suggesting a better concordance among younger groups.[5] However, claims for patients age ≥ 65 years in this study may be incomplete because only those with commercial plans at some time after cancer diagnosis were included. Another study compared RT documented in two large SEER sites (Los Angeles and Detroit) with self-report through a survey of patients age 21-79 years.[27] Using self-report as the gold standard, risk of unascertainment (defined as self-reported RT not documented in SEER) was higher among younger patients (age < 50 years) compared

FIG 2. Absolute difference in radiation starting dates (in days) between ACR and APCD claims. ACR, Arkansas Cancer Registry; APCD, All-Payer Claims Database.
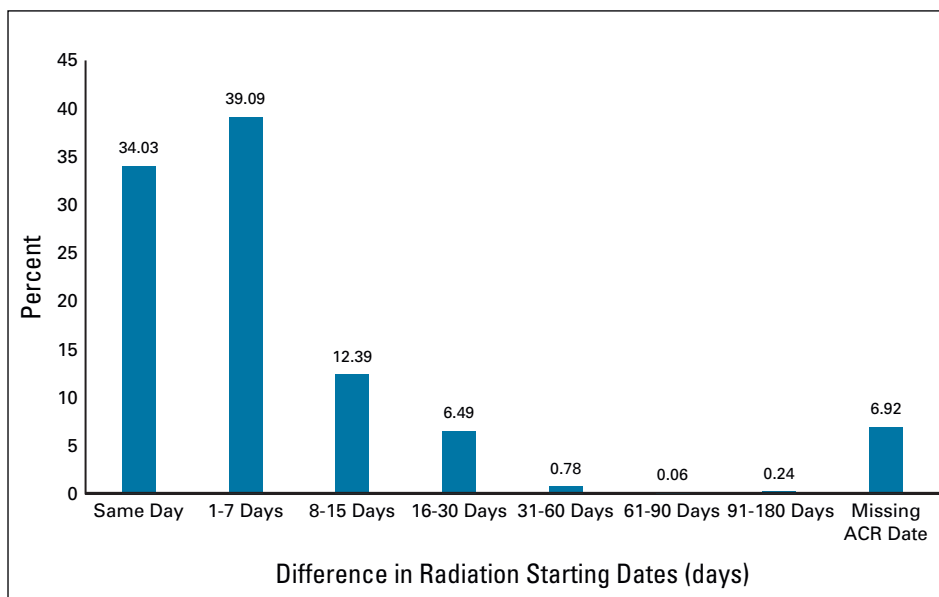
TABLE 4. Logistic Regression of Factors Associated With Disagreement Between Arkansas Cancer Registry and All-Payer Claims Database

Variable	OR	95% CI	P
Age group (ref: ≥ 66 years), years			
18-63	1.56	1.15 to 2.11	.009
64-65	1.60	1.07 to 2.40	.038
Race (ref = non-Hispanic White)			
Hispanic	1.06	0.47 to 2.41	.854
Non-Hispanic Black	1.25	0.92 to 1.70	.137
Non-Hispanic Others	1.86	0.86 to 4.03	.081
Unknown[a]	3.33	0.61 to 18.33	.175
Rural/urban (ref = urban)			
Rural	0.90	0.73 to 1.10	.343
Missing[a]	2.76	0.30 to 25.59	.352
Stage (ref = 0)			
1	1.38	1.03 to 1.85	.044
2	1.25	0.90 to 1.73	.376
3	1.16	0.61 to 2.19	.711
Insurance group (ref = Medicare FFS only)			
Private only	2.45	1.73 to 3.47	< .0001
Medicaid only	1.49	0.76 to 2.90	.220
Private and Medicaid	1.41	0.89 to 2.22	.067
Medicare and private	1.44	1.00 to 2.08	.072
Medicare and Medicaid	1.02	0.56 to 1.84	.666
All 3	1.46	0.70 to 3.07	.829
Year of diagnosis (ref = 2013)			
2014	0.96	0.68 to 1.35	.755
2015	0.90	0.64 to 1.26	.516
2016	1.07	0.77 to 1.49	.799
2017	0.85	0.60 to 1.19	.248

Abbreviations: FFS, fee-for-service; OR, odds ratio.
[a]Estimate may not be reliable because of small sample sizes (< 10).

with those ≥ 65 years (Los Angeles: 43% v 23%; Detroit: 14% v 9%).[27] We also found better agreement in patients with stage 0 cancer compared with those with more advanced stages. This is consistent with previous studies[16,19] and has been attributed to more complex treatment plans for advanced disease that could delay RT.

Assessing time to treatment is an important aspect of cancer care delivery. Most published studies relied on claims database to assess time to RT initiation.[28-31] Some used RT starting date in cancer registry.[32,33] However, no studies have compared the documented starting date between cancer registries and claims. In this study, we found high consistency in RT starting dates between the two sources. Although only 34% had first RT claim on the same day as the RT starting date in ACR, the average gap between the two sources was only 2 days. About 7% of patients with documented RT in both sources had missing RT starting date in ACR, which can be supplemented with information from APCD. This study shows that both data sources can be used to reliably document RT initiation when only one data source is available. APCD can supplement ACR when dates are missing in ACR.

Similar to our study, most published studies used CPT/HCPCS codes in outpatient settings to document RT in claims database.[5,15,23,24] Some also used ICD diagnosis codes for general radiation oncology services,[16] revenue codes,[16,19,25,26] and ICD procedure codes,[16,19,26] which are used to document procedures performed in the inpatient setting. In our cohort, adding those codes only identified 27 additional patients in Medicare inpatient claims. However, these codes are not specific to any anatomic sites. Any delivered RT should have a physician claim with a radiation-related CPT/HCPCS code. Clinically, it is also very unlikely that RT aimed to treat early-stage BC is delivered in an inpatient setting. Thus, the nature of these procedures identified only by ICD diagnosis codes, revenue codes, and/or ICD procedure codes is unclear and including claims identified solely by these codes increased the gap between the starting dates documented in ACR and APCD (average 6 days). Therefore, we recommend to not use these codes to identify RT receipt.

A major limitation is that data were from one state. However, given that an increasing number of states developed or are developing their APCD and potentially link APCD to state's cancer registry, our findings have important implications and could be compared in the future with other states when such data become available. Findings from this study may not generalize to patients with breast cancer excluded from this study: patients not in the APCD, did not receive surgery, HR− or with unknown HR status, stage 4 cancer or with unknown stage, or had multiple cancers. Findings also may not be generalized to small private plans with < 2,000 enrollees or self-insured plans who are not mandated to submit claims to APCD and care provided at the Veterans Health Administration. Since all patients had continuous insurance coverage during the 13 months on/after BC diagnosis, findings could not be generalized to those with interruptions in coverage, who were uninsured, or who received uncompensated care during this time.

In conclusion, in this study, to our knowledge, we conducted the first study comparing RT in Arkansas's APCD linked with state's cancer registry. Overall, we found good consistency between the sources with high sensitivity and PPV and moderate specificity and NPV. Medicare FFS claims had the best consistency with the cancer registry, whereas private claims, which included claims from multiple private payers, had the worse consistency. Consistency was lower among younger patients, private-insured patients, or those with more advantaged disease.

AFFILIATIONS

[1]Division of Pharmaceutical Evaluation and Policy, Department of Pharmacy Practice, College of Pharmacy, University of Arkansas for Medical Sciences, Little Rock, AR

[2]Division of Hematology and Oncology, Department of Internal Medicine, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR

[3]Cancer Administration Service Line, Cancer Registry, University of Arkansas for Medical Sciences, Little Rock, AR

CORRESPONDING AUTHOR

Chenghui Li, PhD, Division of Pharmaceutical Evaluation and Policy, Department of Pharmacy Practice, College of Pharmacy, University of Arkansas for Medical Sciences, 4301 West Markham St, Little Rock, AR 72205; e-mail: CLi@uams.edu.

AUTHOR CONTRIBUTIONS

Conception and design: Chenghui Li, Laura F. Hutchins
Collection and assembly of data: Chenghui Li, Laura F. Hutchins
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

REFERENCES

1. All-Payer Claims Databases. Rockville, MD, Agency for Healthcare Research and Quality. https://www.ahrq.gov/data/apcd/index.htmlhttps://www.ahrq.gov/data/apcd/index.html

2. Introduction: Inventory and Prioritization of Measures to Support the Growing Effort in Transparency Using All-Payer Claims Databases. Rockville, MD, Agency for Healthcare Research and Quality. https://www.ahrq.gov/data/apcd/backgroundrpt/intro.html

3. McCarthy D: State All-Payer Claims Databases: Tools for Improving Health Care Value, Part 1 — How States Establish an APCD and Make It Functional. The Commonwealth Fund Report, 2020

4. Reviewing the Landscape of All-Payer Claims Databases. Rockville, MD, Agency for Healthcare Research and Quality. https://www.ahrq.gov/data/apcd/backgroundrpt/review.html

5. Hashibe M, Ou JY, Herget K, et al: Feasibility of capturing cancer treatment data in the Utah All-Payer Claims Database. JCO Clin Cancer Inform 3:1-10, 2019

6. Fisher B, Redmond C, Poisson R, et al: Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 320:822-828, 1989

7. Veronesi U, Cascinelli N, Mariani L, et al: Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. N Engl J Med 347:1227-1232, 2002

8. Wang J, Deng JP, Sun JY, et al: Noninferior outcome after breast-conserving treatment compared to mastectomy in breast cancer patients with four or more positive lymph nodes. Front Oncol. 9:143, 2019

9. Vila J, Gandini S, Gentilini O: Overall survival according to type of surgery in young (≤40 years) early breast cancer patients: A systematic meta-analysis comparing breast-conserving surgery versus mastectomy. Breast. 24:175-181, 2015

10. Ye JC, Yan W, Christos PJ, et al: Equivalent survival with mastectomy or breast-conserving surgery plus radiation in young women aged < 40 years with early-stage breast cancer: A national registry-based stage-by-stage comparison. Clin Breast Cancer. 15:390-397, 2015

11. Salerno KE: NCCN guidelines update: Evolving radiation therapy recommendations for breast cancer. J Natl Compr Canc Netw. 15:682-684, 2017

12. Buchholz TA, Theriault RL, Niland JC, et al: The use of radiation as a component of breast conservation therapy in National Comprehensive Cancer Network Centers. J Clin Oncol 24:361-369, 2006

13. ACHI: Arkansas Center for Health Improvement. https://achi.net/

14. Money K: Using APCD Unique IDs for Data Linkage. Arkansas Center for Health Improvement, 2018. https://www.nahdo.org/sites/default/files/Kenley%20Money%20%20NAHDO%20AR%20APCD%20Unique%20id%20presentation%2020181003%20FINAL.pdf

15. Walker GV, Giordano SH, Williams M, et al: Muddy water? Variation in reporting receipt of breast cancer radiation therapy by population-based tumor registries. Int J Radiat Oncol Biol Phys 86:686-693, 2013

16. Noone AM, Lund JL, Mariotto A, et al: Comparison of SEER treatment data with Medicare claims. Med Care 54:e55-e64, 2016

17. Facility Oncology Registry Data Standard (FORDS) for 2016. American College of Surgeons, 2016. https://www.facs.org/media/r5rl5scw/fords-2016.pdf

18. Ingram DD, Franco SJ: 2013 NCHS urban–rural classification scheme for counties. National Center for Health Statistics. Vital Health Stat 2:1-73, 2014.

19. 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. J Clin Epidemiol 52:463-470, 1999

20. Cohen JA: Coefficient of agreement for nominal scales. Educ Psychol Meas 20:37-46, 1960

21. McHugh ML: Interrater reliability: The kappa statistic. Biochem Med (Zagreb) 22:276-282, 2012

22. Perraillon MC, Liang R, Sabik LM, et al: The role of all-payer claims databases to expand central cancer registries: Experience from Colorado. Health Serv Res 57:703-711, 2022

23. Doebbeling BN, Wyant DK, McCoy KD, et al: Linked insurance-tumor registry database for health services research. Med Care 37:1105-1115, 1999

24. Hillner BE, McDonald MK, Penberthy L, et al: Measuring standards of care for early breast cancer in an insured population. J Clin Oncol 15:1401-1408, 1997

25. Cooper GS, Yuan Z, Stange KC, et al: Agreement of Medicare claims and tumor registry data for assessment of cancer-related treatment. Med Care 38:411-421, 2000

26. Virnig BA, Warren JL, Cooper GS, et al: Studying radiation therapy using SEER-Medicare-linked data. Med Care 40:IV-49-IV-54, 2002 (8 suppl)

27. Jagsi R, Abrahamse P, Hawley ST, et al: Underascertainment of radiotherapy receipt in Surveillance, Epidemiology, and End Results registry data. Cancer 118:333-341, 2012

28. Gomez DR, Liao KP, Swisher SG, et al: Time to treatment as a quality metric in lung cancer: Staging studies, time to treatment, and patient survival. Radiother Oncol 115:257-263, 2015

29. Nadpara P, Madhavan SS, Tworek C: Guideline-concordant timely lung cancer care and prognosis among elderly patients in the United States: A population-based study. Cancer Epidemiol 39:1136-1144, 2015

30. Onukwugha E, Yong C, Naslund M, et al: Specialist visits and initiation of cancer-directed treatment among a large cohort of men diagnosed with prostate cancer. Urol Oncol 35:150.e17-150.e23, 2017

31. Stokes WA, Hendrix LH, Royce TJ, et al: Racial differences in time from prostate cancer diagnosis to treatment initiation: A population-based study. Cancer 119:2486-2493, 2013

32. Jabo B, Lin AC, Aljehani MA, et al: Impact of breast reconstruction on time to definitive surgical treatment, adjuvant therapy, and breast cancer outcomes. Ann Surg Oncol 25:3096-3105, 2018

33. Bui A, Yang L, Myint A, et al: Race, ethnicity, and socioeconomic status are associated with prolonged time to treatment after a diagnosis of colorectal cancer: A large population-based study. Gastroenterology 160:1394-1396.e3, 2021

APPENDIX

TABLE A1. Procedure Codes for Radiation Therapy Delivery

Code Type	Code
CPT	76950
CPT	77014
CPT	77371
CPT	77372
CPT	77373
CPT	77385
CPT	77386
CPT	77387
CPT	77401
CPT	77402
CPT	77403
CPT	77404
CPT	77406
CPT	77407
CPT	77408
CPT	77409
CPT	77411
CPT	77412
CPT	77413
CPT	77414
CPT	77416
CPT	77417
CPT	77418
CPT	77421
CPT	77422
CPT	77423
CPT	77424
CPT	77425
CPT	77520
CPT	77521
CPT	77522
CPT	77523
CPT	77524
CPT	77525
CPT	77600
CPT	77601
CPT	77602
CPT	77603
CPT	77604
CPT	77605
CPT	77606
CPT	77607

(Continued in next column)

TABLE A1. Procedure Codes for Radiation Therapy Delivery (Continued)

Code Type	Code
CPT	77608
CPT	77609
CPT	77610
CPT	77611
CPT	77612
CPT	77613
CPT	77614
CPT	77615
CPT	77616
CPT	77617
CPT	77618
CPT	77619
CPT	77620
CPT	77750
CPT	77761
CPT	77762
CPT	77763
CPT	77770
CPT	77771
CPT	77772
CPT	77778
CPT	77789
CPT	77790
CPT	0073T
CPT	0197T
CPT	0394T
CPT	0394T
CPT	0395T
CPT	0395T
HCPCS	G6001
HCPCS	G6002
HCPCS	G6003
HCPCS	G6004
HCPCS	G6005
HCPCS	G6006
HCPCS	G6007
HCPCS	G6008
HCPCS	G6009
HCPCS	G6010
HCPCS	G6011
HCPCS	G6012

(Continued on following page)

TABLE A1. Procedure Codes for Radiation Therapy Delivery (Continued)

Code Type	Code
HCPCS	G6013
HCPCS	G6014
HCPCS	G6015
HCPCS	G6016
HCPCS	G6017

TABLE A2. Previously Published Studies Assessing Agreement in Receipt of Radiation Therapy Between Cancer Registry and Claims Database

Article	Year of Diagnosis	Registry	Comparison	Cancer Site	Stage at Diagnosis	Age, Years	Surgery	Follow-Up Window for Identifying Radiation Therapy	Codes Used for Radiation Therapy	Using Claims as the Gold Standard					
										Sensitivity, %	Specificity, %	PPV, %	NPV, %	Agreement, %	Kappa
Cooper et al[25]	1984-1993	10 SEER sites	Medicare	6 sites/breast separately reported	Excluded in situ carcinoma and included unstaged	≥ 65	NA	Within 4 months before or after diagnosis; sensitivity analysis using 6 months	CPT/HCPCS, revenue center codes	57.5	92.7	71.3	87.4	84.3	0.54
Du et al[19]	1992	8 SEER sites	Medicare	Breast	All stages, including unstaged	65-74	Mastectomy, BCS, no surgery	Within 4 months after initial therapy (radiation and/or surgery given within 4 months after diagnosis)	CPT/HCPCS, ICD9 procedure codes, revenue center codes	81.3	96.4	92.6	90.2	91.0	0.80
Virnig et al[26]	1991-1996	9 SEER sites	Medicare	5 sites/breast separately reported	NA	≥ 65	NA	Within 9 months of diagnosis	CPT/HCPCS, ICD9 procedure codes, revenue center codes	86.5	98.2	96.2	93.1	94.2	0.87
Walker et al[15]	2001-2007	16 SEER sites and 3 non-SEER sites (FL, TX, NY)	Medicare	Breast	Stages 1-3	≥ 66	Mastectomy, BCS, biopsy	Within 1 year of diagnosis	CPT/HCPCS	81.30	97.10	84.20	97.10	89.3	0.79
Noone et al[16]	2000-2006	17 SEER sites	Medicare	6 sites/female breast separately reported	All stages, including unstaged	≥ 65	NA	Within 1 year of diagnosis; sensitivity analysis using 4 and 8 months	CPT/HCPCS, ICD9 procedure codes, ICD9 diagnosis codes, revenue center codes	79.6	97.8	97.5	81.8	88.4	0.77
Hillner et al[24]	1989-1991	Virginia state registry	Blue Cross/Blue Shield of Virginia	Breast	Local or regional	< 65	BCS, lumpectomy, nodal dissection	Any time after surgery	CPT/HCPCS	89.3	93.3	82.8	96.0	92.3	0.81
Doebbeling et al[23]	1994	1 SEER site (IA)	Blue Cross/Blue Shield of Iowa	Breast	All stages, including unstaged	21-64	Mastectomy, partial mastectomy, excision of cyst, fibrial and adenoma, or other benign or malignant tumor	NA	CPT/HCPCS	NA	NA	NA	NA	87	NA
Hashibe et al[5]	1/2013-6/2014	1 SEER site (UT)	APCD commercial plans and medical record abstraction	All sites/breast no separately reported	All stages, including unstaged	All ages	NA	Within 1 year of diagnosis	CPT/HCPCS	85.9	89.3	70.1	95.6	88.6	0.70

Abbreviations: APCD, All-Payer Claims Database; BCS, breast-conserving surgery; NPV, negative predictive value; PPV, positive predictive value.