Feasibility of Capturing Cancer Treatment Data in the Utah All-Payer Claims Database

Mia Hashibe, PhD^{1,2}; Judy Y. Ou, PhD^{1,2}; Kimberly Herget, MStat¹; Dan Bolton, PhD²; Jordan McPherson, MS, PharmD²; Charles Hawley, MA³; Jennifer Garvin, MBA, PhD⁴; Jennifer A. Doherty, MS, PhD^{1,2}; and Carol Sweeney, PhD^{1,2}

PURPOSE Incompleteness of treatment data is a recognized limitation of cancer registry data. An all-payer claims database (APCD) is a tool that states use to capture health care information across systems and payer. We linked the Utah Cancer Registry (UCR) records to Utah's statewide APCD and evaluated how this linkage led to improvements in the capture of cancer treatment information.

METHODS We linked cancers diagnosed and reported to the UCR with Utah APCD claims for the calendar years 2013 and 2014 using LinkPlus Software. For patients with breast or colorectal cancers, manual abstraction was completed to provide a gold-standard comparison for the treatment data obtained from the claims.

RESULTS Among 10,759 reportable cancer occurrences linked to the APCD, the claims identified additional patients with cancer who received therapies that had been unknown to the registry, increasing the proportion treated with chemotherapy from 23.7% to 27.6%, hormone therapy from 14.1% to 18.8%, immunotherapy from 4.3% to 13.2%, and radiation therapy from 24.9% to 27.5%. The APCD increased the sensitivity of treatment variables compared with the abstraction gold standard. Notably, sensitivity of hormonal therapy for breast cancer increased from 78.6% to 95.2% when augmented with APCD claims data. However, the APCD alone did not achieve as high specificity for treatment data as did the data collected through traditional registry methods.

CONCLUSIONS This is the first study, to our knowledge, showing that linking cancer registry data with a statewide claims database that covers multiple insurance companies improves cancer treatment data collection. Linking of cancer registry and APCD data can improve comprehensiveness of cancer registry treatment data.

JCO Clin Cancer Inform. © 2019 by American Society of Clinical Oncology

INTRODUCTION

Chemotherapy, hormonal therapy, and immunotherapy data collected by central cancer registries are not made available in the National Cancer Institute (NCI) SEER research data due to concerns about the documented under-reporting of these treatments.^{1,2} As the number of new chemotherapy agents and targeted drugs approved for cancer treatment increases, estimating the population-based survival differences related to these treatments is increasingly important. Survival rates from clinical trials are available for specific new treatments but usually include a restrictive study population with extensive eligibility criteria. Thus, population-based estimates for survival by cancer treatment would be beneficial more broadly for patients with cancer and would offer greater representation. Additionally, because the number of cancer survivors is increasing,³ the long-term effects of these treatments must be assessed with population-based approaches. Previous research using linkage of cancer registry cases to Medicare claims showed that claims data can capture cancer treatments that are

under-reported to central cancer registries, but that results were restricted to patients age ≥ 65 years.²

An all-payer claims database (APCD) is a tool used by states to understand health care use by aggregating claims for all providers across all health care insurance payers.⁴ APCDs can provide claims information for cancer treatments provided in health care settings, including hospitals, outpatient clinics, private practice, and pharmacies. Thus, APCDs are an important potential resource to overcome the issue of under-reporting cancer treatment in registry surveillance data.⁴ Multiple states have APCDs in various stages of development.⁴ As of July 2019, there were 14 states with existing APCDs, four states with APCDs in the implementation phase, and nine states with strong interest.⁴ The Utah Department of Health is one of the early developers and adopters of an APCD.⁵ Utah's APCD is estimated to represent more than 90% of those covered by commercial health insurance plans from the group and individual market.⁶ To our knowledge, no prior reports have evaluated the utility

ASSOCIATED CONTENT

Appendix

Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on October 9, 2019 and published at ascopubs.org/journal/cci on November 22, 2019; DOI <https://doi.org/10.1200/CCI.19.00027>

CONTEXT

Key Objective

To evaluate whether the all-payer claims database (APCD) in Utah can improve the cancer treatment data in the Utah Cancer Registry. This is the first study, to our knowledge, investigating whether a statewide claims database that covers multiple insurance companies can improve cancer treatment data collection.

Knowledge Generated

The APCD improved cancer treatment data for chemotherapy, hormone therapy, immunotherapy, and radiation therapy. Linking of cancer registry and APCD data can improve comprehensiveness of cancer registry treatment data.

Relevance

Improved cancer treatment data in population-based cancer registries can be used to provide population-based survival according to specific treatments in a population of patients with cancer that is broader and more representative of patients.

of a statewide APCD to augment central cancer registry treatment data.

The Utah Cancer Registry (UCR) is one of the original NCI SEER registries, joining SEER in 1973. The UCR covers a population of approximately 2.8 million individuals, according to the 2010 census. The Utah population is 86.1% white, 6.0% other (reporting multiple races/ethnicities), 2.7% two or more races/ethnicities, 2.0% Asian, 1.2% American Indian/Alaskan native, 1.1% black, and 0.9% native Hawaiian or other Pacific Islander. Approximately 13.0% of the Utah population is Hispanic. The UCR linked records for 82% of individual patients with cancer with the Utah APCD, with varying success by cancer diagnosis age, cancer site, and insurance type, as previously reported.⁷ Our aim in this report is to investigate the improvement in registry variables representing the first course of cancer-directed therapy through linkage to APCD.

METHODS

Patients

We identified Utah residents diagnosed with a first primary cancer between January 1, 2013, and June 30, 2014. Eligibility criteria were that the patient with cancer was (1) a Utah resident diagnosed with a SEER-reportable cancer and (2) diagnosed with a first primary cancer and no subsequent primaries. Excluded were patients whose cancers were identified by death certificate only or autopsy only as well as any cancer occurrences only reported to UCR by the Veterans Health Administration. Patients identified by death certificate only or autopsy only are unlikely to have any cancer treatment data. Because data from the Veterans Health Administration are not available in the APCD, we did not include patients reported by that organization. We included patients with cancer of all ages, reportable behaviors, and cancer sites. This study was approved by the University of Utah institutional review board.

Additional details on the APCD are available in the Appendix. We used cancer registry treatment variables for each patient for first course therapy. We created APCD

treatment variables using codes from the Current Procedural Terminology (CPT), Healthcare Common Procedure Coding System (HCPCS), and National Drug Code Directory to indicate if the treatment was found in the APCD. Codes corresponding to cancer therapies were compiled from lists used by prior investigations using claims to study cancer treatment information⁷ and from a review of CPT manuals for new codes. The combined list, provided as a table in the Data Supplement, is a combination of multiple data sources, including SEER*RX, the Cancer Medications Inquiry Database, literature review, and manual review of HCPCS/CPT books for codes.⁸⁻¹⁷ The Cancer Medications Inquiry Database is available in the SEER Oncology Toolbox.¹⁸ Because the intent was to evaluate APCD as a source to augment registry variables representing the first course of therapy, only claims for treatment during the time window from cancer diagnosis to 1 year after the date of diagnosis were used to determine cancer treatment. The 1-year window was selected to be consistent with the SEER program Coding and Staging Manual used by cancer registries and with previous studies of SEER Medicare.^{1,2}

We set a gold standard for validation of treatment obtained through APCD linkage by manual abstraction of medical records. Medical record abstraction was completed for a small subsample of data from eligible patients with breast or colorectal cancer who linked to the APCD, who had a sequence of 00 (first and only cancer diagnosis), and ≤ 64 years of age at diagnosis. We originally expected that data among older patients (ie, ≥ 65 years) would not be as available, because the majority of them would be enrolled in Medicare, but we have since discovered that the APCD is useful for the older population as well. The original design of the study had assumed poor data for older patients. The abstraction was carried out using an abstraction form based on past Patterns of Care¹⁹ study abstraction forms. The abstraction was carried out by a trained certified tumor registrar (CTR). The Patterns of Care abstractions are completed on select cancers for a small subset of the patients each year by the SEER registries. The abstraction

forms include chemotherapy agents, radiation therapy dose, and recurrence, and the form data are more detailed than the treatment data abstracted by the CTR for the core cancer registry data. This abstraction was conducted using electronic medical records for 389 (98%) of 397 patients. The remaining records were abstracted using paper files in physician's offices. We used manual abstraction as the gold standard because the CTRs are experienced in identifying cancer treatment data as part of standard registry operations and for projects such as the NCI Patterns of Care study.¹⁹

Data Analysis

We created indicator variables for each UCR case with regard to treatment received according to UCR records alone and to indicate if cancer treatment was identified in the APCD. By contrasting these variables, we reported the number and percentage of patients with cancer with any indication of cancer treatment before and after linkage. For the comparison of cancer treatment variables using both cancer registry and APCD, to a gold standard of manual abstraction, we estimated sensitivity and specificity for patients with breast or colorectal cancer only. We identified characteristics of groups for whom the linkage improved a treatment report using logistic regression models to estimate odds ratios and 95% confidence levels.

RESULTS

The number of eligible patients with cancer for this project, the proportion linked, and the proportion with new

treatment data identified in the APCD are shown in Figure 1. Of the 13,533 reportable cancer occurrences, 7,609 patients were < 65 years old at cancer diagnosis, and 5,924 were ≥ 65 years at cancer diagnosis (Table 1). Among these patients, 1,358 (17.8%) of 7,609 who were < 65 years old and 896 (15.1%) of 5,924 who were ≥ 65 years had additional treatment data observed in the Utah APCD that was not observed in the registry data alone. Of the patients who had additional treatment data from the APCD, there was a higher proportion of women in both age groups compared with men, more black patients compared with other race/ethnicity groups, and more urban patients compared with rural patients.

Of the 13,533 reportable cancer occurrences, 10,759 patients with cancer (79.5%) diagnosed between January 2013 and June 2014 had data linked to the APCD for claims filed between January 1, 2013 and December 31, 2014. Among the 10,759 patients, claims for chemotherapy were identified in 24.1% (Table 2). Of the patients marked as not having cancer treatment in the cancer registry data, the APCD identified an additional 497 patients who had received chemotherapy, 590 patients treated with hormone therapy, 326 patients treated with radiation therapy, and 1,190 patients treated with immunotherapy. The APCD also provides the added value of the chemotherapy agent names and the patients' duration of treatment.

The overall proportion of patients with cancer who received cancer treatment increased when registry data were augmented with APCD (Fig 2). The claims identified additional patients who received therapies that had been unknown to the registry, increasing the proportion treated with chemotherapy from 23.7% to 27.6%, hormone therapy from 14.1% to 18.8%, immunotherapy from 4.3% to 13.2%, and radiation therapy from 24.9% to 27.5%. However, there were also appreciable proportions of patients with the reverse (ie, therapy identified through traditional registry data collection but not from APCD). The contribution of APCD alone was prominent for hormone therapy among patients with breast cancer, for chemotherapy among patients with lung or colorectal cancers, and for radiation therapy among patients with lung or prostate cancers. Overall, the contribution to melanoma treatment was largely for biologic therapy and from the APCD alone.

The improvements to sensitivity and specificity after augmenting cancer registry treatment variables with the APCD are listed in Table 3 and in Appendix Figure A1 for breast and colorectal cancers, with the gold standard of manual abstraction by a trained CTR. Even after augmenting with APCD, there was an indication of under-reporting of chemotherapy for breast cancer by the cancer registry variable compared with abstraction. A factor contributing to cancers that were coded as treated through the augmented cancer registry variable, but not from abstraction, was when the therapy was determined by the abstractor to not be the first

FIG 1. Overall schema for patients with cancer linked to the all-payer claims data (APCD).

TABLE 1. Additional Treatment Information Obtained From Utah APCD Among All Utah Patients With Cancer Diagnosed Between January 2013 and June 2014 by Demographic and Diagnostic Characteristics

Characteristic	< 65 Years			≥ 65 Years		
	Total No.	Any New Treatment Data From APCD		Total No.	Any New Treatment Data From APCD	
		No.	%		No.	%
Total	7,609	1,358	17.8	5,924	896	15.1
Age at diagnosis, years						
< 50	3,090	548	17.7	—	—	—
50-64	4,519	810	17.9	—	—	—
65-74	—	—	—	3,318	515	15.5
≥ 75	—	—	—	2,606	381	14.6
Sex						
Male	3,412	494	14.5	3,218	435	13.5
Female	4,197	864	20.6	2,706	461	17.0
Year of diagnosis						
2013	5,011	942	18.8	3,906	609	15.6
2014	2,598	416	16.0	2,018	287	14.2
Race/ethnicity						
Non-Hispanic white	6,458	1,211	18.8	5,340	823	15.4
Hispanic, any race	607	81	13.3	247	43	17.4
Black	83	21	25.3	34	7	20.6
American Indian or Alaskan native	36	5	13.9	27	2	7.4
Asian or Pacific Islander	183	29	15.8	95	10	10.5
Unknown	242	11	4.5	181	11	6.1
County of residence						
Urban	6,735	1,212	18.0	5,114	794	15.5
Rural	874	146	16.7	810	102	12.6
Site						
Colorectal	495	95	19.2	424	37	8.7
Breast	1,237	374	30.2	791	202	25.5
Lung	223	52	23.3	447	61	13.6
Melanoma	1,061	104	9.8	684	68	9.9
Other	3,875	656	16.9	2,569	389	15.1
Prostate	718	77	10.7	1,009	139	13.8
Primary payer at diagnosis						
Private insurance	5,058	1,047	20.7	528	130	24.6
Medicaid	438	118	26.9	54	12	22.2
Medicare	265	53	20.0	3,114	504	16.2
Medicare and Medicaid, NOS	105	8	7.6	801	133	16.6
Not insured	417	46	11.0	46	3	6.5

(Continued in next column)

TABLE 1. Additional Treatment Information Obtained From Utah APCD Among All Utah Patients With Cancer Diagnosed Between January 2013 and June 2014 by Demographic and Diagnostic Characteristics (Continued)

Characteristic	< 65 Years			≥ 65 Years		
	Total No.	Any New Treatment Data From APCD		Total No.	Any New Treatment Data From APCD	
		No.	%		No.	%
Other government*	283	10	3.5	251	8	3.2
Unknown	1,043	76	7.3	1,130	106	9.4
Vital status						
Alive	6,768	1,138	16.8	4,208	663	15.8
Deceased	281	29	10.3	822	52	6.3
< 3 months						
Deceased	122	22	18.0	248	29	11.7
4-6 months						
Deceased	174	54	31.0	262	46	17.6
7-11 months						
Deceased ≥ 1 year	264	115	43.6	384	106	27.6
Reporting source						
Inpatient or hospital	6,701	1,308	19.5	4,846	805	16.6
Outpatient†	908	50	5.5	1,078	91	8.4
Behavior						
Benign	468	31	6.6	247	22	8.9
Borderline	38	4	10.5	20	3	15.0
In situ	833	82	9.8	646	68	10.5
Invasive	6,270	1,241	19.8	5,011	803	16.0

NOTE. All row variables are cancer registry variables.

Abbreviations: APCD, all-payer claims data; NOS, not otherwise specified.

*Includes Veterans Health Administration, TRICARE, and Indian Health Services.

†Includes radiation clinics, pathology labs, physician clinics, nursing homes and hospice.

course. For patients with colorectal cancer, augmenting with APCD increased the sensitivity from 88.1% to 99.0% for chemotherapy and from 91.1% to 95.6% for radiation therapy. The specificity, conversely, appeared to decrease fairly consistently for both colorectal cancer and breast cancer.

The likelihood of finding cancer-directed treatment in the APCD did not differ greatly by race/ethnicity or rural/urban residence (Table 4). There appeared to be higher odds of identifying chemotherapy for younger patients compared with older patients and for local stage compared with distant stage. For hormone therapy, there were higher odds of identifying treatment of in situ cancers compared with local-stage cancers and lower odds of identifying treatment of younger patients compared with older patients. Patients

TABLE 2. Comparison Between the Cancer Registry and APCD of First Course of Cancer Treatment Among Utah Patients With Cancer Diagnosed Between January 2013 and June 2014 and Linked to the APCD

Cancer Registry Variable by Treatment		Total No. of Patients		Claim for Therapy Found in APCD	
				Yes	No
		No.	%	No.	%
Chemotherapy					
Total	10,759	2,592	24.1	8,167	75.9
Yes	2,714	2,059	75.9	655	24.1
NOS	171	131	76.6	40	23.4
Single agent	676	443	65.5	233	34.5
Multiple agents	1,867	1,485	79.5	382	20.5
No	7,588	497	6.5	7,091	93.5
Unknown	457	36	7.9	421	92.1
Hormone					
Total	10,759	1,506	14.0	9,253	86.0
Yes	1,625	878	54.0	747	46.0
No	8,678	590	6.8	8,088	93.2
Unknown	456	38	8.3	418	91.7
Radiation					
Total	10,759	2,348	21.8	8,411	78.2
Yes	2,845	1,994	70.1	851	29.9
Beam radiation	2,067	1,573	76.1	494	23.9
Radioactive implants	357	271	75.9	86	24.1
Radioisotopes	262	27	10.3	235	89.7
Combination*	126	96	76.2	30	23.8
Radiation, NOS	33	27	81.8	6	18.2
No	7,456	326	4.4	7,130	95.6
Unknown	458	28	6.1	430	93.9
Biologic therapy					
Total	10,759	1,478	13.7	9,281	86.3
Yes	468	273	58.3	195	41.7
No	9,935	1,190	12.0	8,745	88.0
Unknown	356	12	3.4	340	96.6

Abbreviations: APCD, all-payer claims data; NOS, not otherwise specified.

*Combination of beam radiation therapy with radioactive implants or radioisotopes.

with cancer who had regional or distant stage were less likely to have hormone therapy identified in the APCD. Radiation treatment was less likely to be identified in the APCD for later stages and for breast, lung, or prostate cancer compared with colorectal cancer. Biologic therapies were more commonly identified in the APCD among patients with breast cancer or melanoma compared with colorectal cancer.

DISCUSSION

This study shows that the APCD is effective in improving treatment data for the Utah cancer registry. The APCD linkage helped to identify hundreds of patients with cancer

who had received chemotherapy, radiation therapy, hormone therapy, or biologic therapies that were not reported to or identified by the registry through other sources. For patients with breast or colorectal cancer, for whom we had a gold-standard comparison of abstracted treatment data by a CTR, we showed that sensitivity improved for most treatments to high sensitivity levels (> 95%) with the APCD. For breast cancer chemotherapy, though the sensitivity improved with APCD as an additional source, it was still fairly low at 69.5%. Additional directions to improve sensitivity and specificity include adjusting windows of time for assessing first-course treatment, depending on the cancer. The APCD was successful in identifying

FIG 2. Proportion of patients with cancer in Utah diagnosed between January 2013 and June 2014 who received chemotherapy, hormone therapy, radiation therapy, and immunotherapy according to data from cancer registry and all-payer claims data (APCD), and by cancer site: (A) all sites; (B) breast; (C) colorectal; (D) lung; (E) melanoma; and (F) prostate.

a range of chemotherapy, hormone therapy, and immunotherapy drugs.

Factors such as age at diagnosis, cancer site, and stage were associated with the likelihood of identifying treatment in the APCD and varied by treatment type. It is unclear why patients with regional or distant-stage disease were less likely to have hormone therapy identified in the APCD. Biologic therapies were more commonly identified among patients with breast cancer or melanoma in the APCD, which may reflect increasing use of immunotherapy, such as immune checkpoint inhibitors, for melanoma and targeted therapies, such as trastuzumab, pertuzumab, and

lapatinib, for breast cancer. Differences observed in finding APCD treatment data thus could reflect both treatment patterns and insurance coverage.

Previous studies have shown a range of improvements in using claims data to improve cancer treatment data in cancer registries. Comparison of registry treatment variables from SEER registries to Medicare treatment data for 433,000 patients with cancer who were ≥ 65 years old showed that the sensitivity of SEER data, using Medicare as the gold standard, for identifying treatments was 68% for chemotherapy, 80% for radiation therapy, and 69% for hormone therapy.² In North Carolina, billing data from three

TABLE 3. Sensitivity and Specificity of Treatment for Breast and Colorectal Cancers, Comparing SEER and APCD to Abstraction As the Gold Standard

Treatment	Breast Cancer		Colorectal Cancer	
	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
Chemotherapy				
SEER only	65.9	97.2	88.1	99.0
APCD only	57.3	88.9	77.2	89.9
Both	69.5	86.1	95.1	88.9
Radiation				
SEER only	95.4	100.0	91.1	100.0
APCD only	74.6	85.7	71.1	97.4
Both	97.7	85.7	95.6	97.4
Hormone				
SEER only	78.6	83.8		
APCD only	74.6	68.9		
Both	95.2	58.1		

Abbreviation: APCD, all-payer claims data.

oncology practices were used to improve chemotherapy and radiation information by 26% and 46%, respectively, for patients with solid tumors.²⁰ Administrative claims data for one payer in Western New York were linked to the National Cancer Database for 439 patients with breast cancer.²¹ That link reported that the proportions of treatment identified by the National Cancer Database compared with the claims data were 38% for radiotherapy, 47% for chemotherapy, and 18% for hormone therapy.²¹ The Iowa Cancer Registry linked data from 4,397 patients with breast cancer (1989-1996) with Blue Cross Blue Shield claims data, and the concordances between the two sources were 79.0% for surgery, 84.0% for chemotherapy, and 87.0% for radiation therapy.²² The Virginia Cancer Registry also linked data to Blue Cross Blue Shield claims data for 1989-1991 for 918 patients with breast cancer.²³ Of the Blue Cross Blue Shield enrollees, the proportion of women who had treatment according to the cancer registry and also had claims for treatments were 94.1% for surgery and 82.7% for radiation therapy. For chemotherapy, 62% more women had claims for chemotherapy compared with the cancer registry chemotherapy data. Our results are unique in using a statewide claims database that covers multiple insurance companies and provide strong evidence that insurance claims databases improve cancer treatment data collection.

Limitations of this study include the less-than-100% coverage of the cancer registry cases by APCD. Patient populations not represented in APCD include those who are not insured, who have their health insurance through small insurance carriers, or who are covered by the Veterans Health Administration or the military TRICARE system—an inherent limitation of the approach. In addition, for the different types of treatment, we observed that the likelihood

of finding cancer-directed treatment in the APCD differed by age at diagnosis, cancer site, and stage. Conversely, we did not observe differences in the likelihood of finding cancer-directed treatment by rural/urban status or race/ethnicity. Future directions of research to improve cancer treatment data include using additional sources of treatment data, such as Medicare data, along with the APCD.

Another limitation is that we restricted the patients to those diagnosed with their first and only primary. The purpose of this eligibility restriction was to curtail the complexity of assessing treatment histories when patients had multiple primaries. However, the lower specificity we observed for the combination of cancer registry and APCD may be due to the complexity of abstracting on the first versus second course of treatment. The low sensitivity we observed for breast cancer chemotherapy even with the combination of cancer registry and APCD data was unexpected. It is unclear why the sensitivity remained low, but additional sources of data clearly are necessary to identify chemotherapy use for patients with breast cancer.

Strengths of this study include the use of an existing data resource to improve cancer treatment data for the cancer registry and the ability to obtain specific chemotherapy agent and hormone therapy names. We confirmed that the APCD improves cancer treatment data collection with a feasible approach based on linking patient data. Biologic therapies were identified more frequently for patients with breast cancer or melanoma than for those with colorectal cancer. Even for patients who were already identified as having had chemotherapy or radiation therapy in the cancer registry, the linkage of APCD to the cancer registry is of potential value for research, because the APCD includes chemotherapy agent names and duration of treatment. We expect that the methods developed in this project can be applied to more recently available years of APCD information.

This study was a pilot study for an initial assessment of the APCD linkage and to determine whether the APCD could improve ascertainment of cancer treatment, both of which have now both been demonstrated successfully. Future work on use of APCD to improve cancer surveillance and on use of a linked cancer registry-APCD in research is warranted. Next steps will include routine linkage for additional years of diagnosis and incorporation of APCD information into registry treatment variables. Claims can be evaluated as a source of data for identifying comorbidities at time of diagnosis of cancer, the late effects of treatment, recurrence after treatment ceases, cancer treatment information for the entire treatment history, and case finding. In conclusion, this pilot study demonstrates the utility of APCD linkage as a cancer treatment data source for cancer registries and research. For states with APCDs available, central cancer registries will benefit from linkage in obtaining more cancer treatment data.

TABLE 4. Patient Characteristics and Likelihood of Finding Cancer-Directed Treatment in the All-Payer Claims Database, Missed by the Utah Cancer Registry

Characteristic	Chemotherapy				Hormone				Radiation				Biologic Therapy			
	Total No.	No. Missed	OR	95% CI	No. Missed	OR	95% CI	No. Missed	OR	95% CI	No. Missed	OR	95% CI	No. Missed	OR	95% CI
Race/ethnicity																
Non-Hispanic white	11,798	489	1.00		564	1.00		323	1.00		1,078	1.00				
Hispanic	854	26	1.37	0.92 to 2.06	36	1.21	0.85 to 1.74	19	1.39	0.87 to 2.23	67	0.81	0.65 to 1.00			
Black	117	< 5	1.67	0.53 to 5.29	< 5	1.40	0.50 to 3.91	< 5	2.12	0.52 to 8.67	23	1.57	0.99 to 2.51			
AIAN	63	< 5	3.02	0.42 to 21.92	< 5	1.49	0.34 to 6.42	< 5	0.86	0.21 to 3.60	< 5	0.73	0.32 to 1.64			
API	278	8	1.47	0.72 to 2.99	10	1.64	0.85 to 3.17	< 5	3.09	0.98 to 9.73	26	1.05	0.74 to 1.48			
Unknown	423	6	2.57	1.13 to 5.86	12	0.79	0.43 to 1.45	5	1.19	0.48 to 2.95	< 5	0.13	0.06 to 0.27			
Age at diagnosis, years																
< 50	3,090	99	1.42	1.08 to 1.88	180	0.53	0.40 to 0.70	76	0.70	0.48 to 1.01	326	1.53	1.29 to 1.81			
50-64	4,519	175	1.11	0.87 to 1.41	216	0.81	0.62 to 1.05	142	0.60	0.43 to 0.83	457	1.66	1.42 to 1.95			
65-74	3,318	141	0.99	0.77 to 1.28	144	0.91	0.69 to 1.21	85	0.78	0.54 to 1.11	248	1.34	1.13 to 1.59			
≥ 75	2,606	118	1.00		88	1.00		51	1.00		174	1.00				
County residence at diagnosis																
Urban	11,849	466	1.00		556	1.00		319	1.00		1,083	1.00				
Rural	1,684	67	1.00	0.77 to 1.30	72	1.04	0.80 to 1.35	35	1.35	0.94 to 1.93	119	0.79	0.66 to 0.93			
Cancer site																
Breast	2,028	81	1.07	0.73 to 1.58	341	0.08	0.05 to 0.13	109	0.25	0.14 to 0.43	257	2.26	1.78 to 2.86			
Colorectal	919	42	1.00		16	1.00		15	1.00		80	1.00				
Lung	670	36	1.03	0.65 to 1.65	0			31	0.47	0.25 to 0.89	71	0.72	0.52 to 0.98			
Melanoma	1,745	58	1.15	0.75 to 1.76	7	3.58	1.45 to 8.83	10	1.98	0.87 to 4.52	108	1.83	1.41 to 2.38			
Prostate	1,727	38	2.11	1.34 to 3.32	120	0.21	0.12 to 0.35	49	0.44	0.24 to 0.79	51	0.34	0.24 to 0.48			
Other	6,444	278	1.06	0.76 to 1.48	144	0.86	0.51 to 1.44	140	0.81	0.47 to 1.39	635	1.14	0.92 to 1.42			
Stage at diagnosis																
In situ	1,449	48	1.12	0.80 to 1.56	36	2.12	1.46 to 3.06	22	1.10	0.68 to 1.78	57	0.76	0.63 to 0.93			
Local	5,978	206	1.00		310	1.00		115	1.00		485	1.00				
Regional	3,096	124	0.93	0.74 to 1.18	184	0.80	0.66 to 0.98	103	0.56	0.43 to 0.74	274	1.28	1.11 to 1.47			
Distant	2,440	133	0.71	0.56 to 0.91	81	0.65	0.50 to 0.86	105	0.34	0.25 to 0.46	367	3.33	2.89 to 3.84			
Unstaged	570	22	0.96	0.60 to 1.51	17	0.92	0.55 to 1.54	9	0.88	0.44 to 1.76	19	0.50	0.32 to 0.77			

NOTE. Includes both linked and unlinked patient data.

Abbreviations: AIAN, American Indian or Alaskan native; API, Asian/Pacific Islander; OR, odds ratio.

AFFILIATIONS

¹Utah Cancer Registry, University of Utah, Salt Lake City, UT

²University of Utah School of Medicine, Salt Lake City, UT

³Utah Department of Health, Salt Lake City, UT

⁴University of Utah, Salt Lake City, UT

CORRESPONDING AUTHOR

Mia Hashibe, PhD, Utah Cancer Registry, Department of Family and Preventive Medicine, University of Utah School of Medicine, Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT 84112; e-mail: mia.hashibe@utah.edu.

SUPPORT

Supported by the National Cancer Institute SEER Program Contract No. RRSS HHSN261201300017/HHSN26100008. The Utah Cancer Registry is funded by the National Cancer Institute SEER Program, Contract No. 75N91018D000016, and the US Center for Disease Control and Prevention National Program of Cancer Registries, Cooperative Agreement No. NU58DP0063200-01, with additional support from the University of Utah and Huntsman Cancer Foundation.

AUTHOR CONTRIBUTIONS

Conception and design: Mia Hashibe, Judy Y. Ou, Kimberly Herget, Jordan McPherson, Jennifer Garvin, Carol Sweeney

Financial support: Mia Hashibe, Jennifer A. Doherty, Carol Sweeney

Administrative support: Charles Hawley, Jennifer A. Doherty

Provision of study material or patients: Charles Hawley, Jennifer Garvin
Collection and assembly of data: Mia Hashibe, Kimberly Herget, Charles Hawley, Carol Sweeney

Data analysis and interpretation: Mia Hashibe, Kimberly Herget, Dan Bolton, Jordan McPherson, Jennifer A. Doherty, Carol Sweeney

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://openpayments.org)).

Jennifer Garvin

Research Funding: Amgen (Inst)

Travel, Accommodations, Expenses: Amgen (Inst)

No other potential conflicts of interest were reported.

REFERENCES

- Warren JL, Harlan LC, Fahey A, et al: Utility of the SEER-Medicare data to identify chemotherapy use. *Med Care* 40:IV-55-IV-61, 2002
- Noone AM, Lund JL, Mariotto A, et al: Comparison of SEER treatment data with Medicare claims. *Med Care* 54:e55-e64, 2016
- American Cancer Society: Cancer Treatment and Survivorship Facts and Figures, 2016-2017. Atlanta, GA, American Cancer Society, 2016.
- All-Payer Claims Database Council: The Basics of All-Payer Claims Databases: A Primer for States. Princeton, NJ, Robert Wood Johnson Foundation, 2014.
- All-Payer Claims Database Council: APCD Council interactive state report map. <http://www.apcdouncil.org/state/map>.
- Utah Insurance Department: 2014 Health Insurance Market Report. <https://insurance.utah.gov/wp-content/uploads/2014HealthMarketReport.pdf>
- Garvin JH, Herget KA, Hashibe M, et al: Linkage between Utah all-payers claims database and central cancer registry. *Health Serv Res* 54:707-713, 2019
- Cooper GS, Virnig B, Klabunde CN, et al: Use of SEER-Medicare data for measuring cancer surgery. *Med Care* 40:IV-43-IV-48, 2002
- Freeman JL, Klabunde CN, Schussler N, et al: Measuring breast, colorectal, and prostate cancer screening with Medicare claims data. *Med Care* 40:IV-36-IV-42, 2002
- Harlan LC, Clegg LX, Trimble EL: Trends in surgery and chemotherapy for women diagnosed with ovarian cancer in the United States. *J Clin Oncol* 21:3488-3494, 2003
- Harlan LC, Clegg LX, Warren JL: Chemotherapy in women with breast cancer. *Ann Intern Med* 139:868, 2003
- Harlan LC, Hankey BF: The Surveillance, Epidemiology, and End Results program database as a resource for conducting descriptive epidemiologic and clinical studies. *J Clin Oncol* 21:2232-2233, 2003
- 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
- Virnig BA, Warren JL, Cooper GS, et al: Studying radiation therapy using SEER-Medicare-linked data. *Med Care* 40:IV-49-IV-54, 2002
- Warren JL, Harlan LC: Can cancer registry data be used to study cancer treatment? *Med Care* 41:1003-1005, 2003
- Warren JL, Harlan LC, Fahey A, et al: Utility of the SEER-Medicare data to identify chemotherapy use. *Med Care* 40:IV-55-IV-61, 2002
- Yu JB, Gross CP, Wilson LD, et al: NCI SEER public-use data: Applications and limitations in oncology research. *Oncology (Williston Park)* 23:288-295, 2009
- National Cancer Institute Surveillance, Epidemiology, and End Results Program: Observational Research in Oncology Toolbox. <https://seer.cancer.gov/oncologytoolbox/>
- Edwards BK, Brown ML, Wingo PA, et al: Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment. *J Natl Cancer Inst* 97:1407-1427, 2005
- Penberthy L, Petkov V, McClish D, et al: The value of billing data from oncology practice to supplement treatment information for cancer surveillance. *J Registry Manag* 41:57-64, 2014
- Meguerditchian AN, Stewart A, Roistacher J, et al: Claims data linked to hospital registry data enhance evaluation of the quality of care of breast cancer. *J Surg Oncol* 101:593-599, 2010
- Doebbeling BN, Wyant DK, McCoy KD, et al: Linked insurance-tumor registry database for health services research. *Med Care* 37:1105-1115, 1999
- 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

APPENDIX

Supplemental Methods

Commercial health insurance carriers with 2,500 or more covered lives in Utah are required to submit to the Utah all-payer claims database (APCD; Utah Department of Health Office of Health Care Statistics: http://stats.health.utah.gov/wp-content/uploads/2016/06/APCD_DSG_v2.2.pdf). Utah's APCD is estimated to represent more than 90% of those covered by group or individual commercial health insurance plans.⁶ Medicaid claims are reported, but Medicare does not submit claims to the state nor is care provided by the Veterans Health Administration or military TRICARE system reported. Claims may be submitted to the APCD by secondary insurance providers for individuals with primary insurance under Medicare (Utah Health Data Committee/Office of Health Care Statistics: http://stats.health.utah.gov/wp-content/uploads/2018/08/APCD_ClaimsCentric_Limited_Use_Datamart_User_Manual_20180814.pdf). Care that is self-paid for by the patient and charity care is not captured by the APCD. The Utah Department of Health makes APCD datasets available for research by request, starting with the year 2013.

The APCD data structure for inpatient and outpatient claims incorporates data elements from the electronic CMS-1500 and UB04

claims forms. Each claim includes identifiers for patient, provider, and insurer, date of service, charges, diagnosis codes, and Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) codes for outpatient encounters, services, and supplies. Inpatient procedure codes are also present using International Classification of Diseases, 10th revision, Procedure Coding System. The pharmacy claims data structure is based on a National Council for Prescription Drug Programs standard format, containing identifiers for patient, provider, and insurer, charges, national drug code, and drug name. Data processing and matching claims for each patient within the APCD for Utah Department of Health was conducted by 3M Health Information Systems (St. Paul, MN) using a deterministic matching algorithm (Link Plus Software, Atlanta, GA; Centers for Disease Control and Prevention: https://www.cdc.gov/cancer/npcr/tools/registryplus/lp_features.htm; Fellegi IP et al: J Am Stat Assoc 64:1183-1210, 1969; Dempster AP et al: J R Stat Soc [Ser A] 39:1-38, 1977). We examined claims filed between January 1, 2013, and December 31, 2014, because this covered the years for which the APCD was available when this project was initiated.

FIG A1. Sensitivity and specificity of cancer registry treatment variables compared with the manual abstraction by certified tumor registrars as the gold standard, considering registry data from traditional sources only, data from all-payer claims data (APCD) claims only, and both—registry augmented with APCD. (A) Breast; (B) colorectal.