Assessment of Oncology Practice Billing Claims for Supplementing Chemotherapy: A Pilot Study in the Georgia SEER Cancer Registry

Kevin Ward, *Emory University*
N Howlader, *National Cancer Institute (NCI)*
JL Warren, *National Cancer Institute (NCI)*
DS Campbell, *Information Management Services, Inc.*
L Coyle, *Information Management Services, Inc.*
AB Mariotto, *National Cancer Institute (NCI)*

Journal Title: Journal of the National Cancer Institute - Monographs

Volume: Volume 2020, Number 55

Publisher: (publisher) | 2020-05-01, Pages 82-88

Type of Work: Article

Publisher DOI: 10.1093/jncimonographs/lgaa006

Permanent URL: https://pid.emory.edu/ark:/25593/vvjmq

Final published version: http://dx.doi.org/10.1093/jncimonographs/lgaa006

OXFORD

J Natl Cancer Inst Monogr (2020) 2020(55): lgaa006

doi: 10.1093/jncimonographs/lgaa006
Article

ARTICLE

Assessment of Oncology Practice Billing Claims for Supplementing Chemotherapy: A Pilot Study in the Georgia SEER Cancer Registry

Nadia Howlader, Kevin C. Ward, Joan L. Warren, Dave S. Campbell, Linda Coyle, Angela B. Mariotto

See the Notes section for the full list of authors' affiliations.
Correspondence to: Nadia Howlader, PhD, MS, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 9609 Medical Center Drive, Bethesda, MD 20892 (e-mail: howladern@mail.nih.gov).

Abstract

Background: Chemotherapy information in the population-based cancer registries is underascertained and lacks detail. We conducted a pilot study in the Georgia SEER Cancer Registry (GCR) to investigate the feasibility of supplementing chemotherapy information using billing claims from six private oncology practices (OP). **Methods:** To assess cancer patients' representativeness from OP, we compared individuals with invasive first primary cancers diagnosed during 2013–2015 in the GCR (cohort 1) with those who had at least one OP claim in the 12 months after diagnosis (cohort 2). To assess completeness of OP claims to capture chemotherapy (yes or no), we further restricted cohort 2 to patients ages 65 years and older enrolled in fee-for-service Medicare Part A and B from the diagnosis date through 12 months follow-up or to the date of death. With Medicare data serving as the gold standard, sensitivity, specificity, and kappa statistics for the receipt of chemotherapy per OP claims were calculated by demographic and clinical characteristics. **Results:** Cancer patients seeking care in the OP included in our analysis were not representative of the underlying patient population in the GCR. The practices underrepresented minorities and uninsured while overrepresenting females, persons with high socioeconomic status, patients residing outside the metropolitan Atlanta area, and persons with advance staged disease. The ability of practice claims to identify chemotherapy receipt was moderate (76.1% sensitivity) but varied by demographic and clinical characteristics (76.1–83.0%). **Conclusions:** Given the limited ability of OP claims to identify chemotherapy receipt, we suggest analyzing these data for hypothesis generation, but inference should be limited to this patient cohort.

The Surveillance, Epidemiology, and End Results (SEER) program, funded by the National Cancer Institute (NCI), is currently composed of 21 population-based cancer registries, covering 34.6% of the total US population and reporting on more than 650 000 incident cancer patients annually (1). SEER registries collect information on first course of cancer treatment, including surgery, chemotherapy, radiation therapy, and hormone therapy.

Hospital medical records have historically been the primary source of information for cancer registries. With more care being delivered in outpatient settings, it is increasingly difficult for cancer registries to capture complete treatment information on all cancer patients diagnosed in their catchment areas. For example, a study that compared receipt of chemotherapy among elderly patients in the SEER data with their Medicare claims found that about 8% of total patients who had a Medicare claim for chemotherapy had no documentation of chemotherapy in the registry data (2). Because of concerns regarding this known underascertainment, NCI does not release information on chemotherapy receipt in the public-use SEER research database.

The increasing availability of electronic health data offers additional potential resources for maximizing the efficiency (eg,

Received: February 12, 2020; **Accepted:** February 12, 2020

quicker, less labor-intensive ascertainment of cancer patients, treatments, and outcomes), comprehensiveness, and completeness of cancer information collected by registries. In turn, more comprehensive, complete data could facilitate expanded data release and enhanced research opportunities. Electronic claims from oncology practices (OP), which include detailed information about cancer patients and the care they receive—including oncolytic agents administered, dates of receipt, and doses—are one type of data that has garnered interest. However, using OP claims to identify patients who receive chemotherapy may be limited because of disparities in healthcare access (3–5). Additionally, even among patients with no barriers to care, OP claims may have limited utility because of a shift in care from private OP to consolidated, hospital-based outpatient clinics. This shift in care setting has resulted in changes in how insurance, most notably the Centers for Medicare and Medicaid Services, reimburses oncologists for the administration of chemotherapy (6). Nonetheless, based on these OP claims, we seek to build an expanded research dataset with all the valuable great clinical details they contain on agents, doses, and dates—similar to the SEER-Medicare linked dataset—that includes younger patients as well as older ones.

Using administrative billing claims from private oncology practices in Georgia, the Georgia Cancer Registry (GCR) conducted a pilot study in collaboration with NCI's SEER program to investigate the feasibility of capturing longitudinal, detailed cancer treatment, with a special focus on chemotherapy. The study's goal was to build a specialized research dataset for cancer research analyses. The first aim was to determine the representativeness of the cancer patients included in OP claims. This is important because cancer registry data are recognized as being population-based, that is, representative of the whole population of their catchment area. The second aim was to determine the completeness of the OP claims to capture chemotherapy among elderly patients in comparison to their Medicare claims. The rationale for this aim is to understand the completeness of chemotherapy information for a specialized research dataset containing patients linked to OP claims. Medicare claims were considered the gold standard because previous studies have indicated that these data are highly accurate for capturing administration of chemotherapy (7–9). The third aim was to study how much is gained by adding the OP claims information to the chemotherapy information (yes or no) already captured by the registry.

Methods

Data Sources

We used individual-level information from three data sources: Georgia SEER Cancer Registry, Georgia oncology practice claims, and SEER-Medicare linked database.

Georgia SEER Data. The GCR collects information on all newly diagnosed cancers within the state. The registry maintains records for each individual in their data, including first and last name, date of birth, and Social Security number. Information obtained by the registry includes the cancer site and stage, month and year of diagnosis, International Classification of Disease-Oncology, third edition, codes, insurance status, demographic characteristics, and reporting source (http://www.seer.cancer.gov). Individual-level socioeconomic status measures are not collected. However, NCI links the SEER registry residential address census tract data to Census Bureau data to provide aggregate socioeconomic measures, including residential census tract median household income, which was categorized into quartiles. The SEER registries routinely collect data on the first course of cancer treatment including information on surgery, radiation therapy, chemotherapy, and hormone therapy. Information on surgery is reported in the publicly available SEER dataset (10,11). NCI does not make data on chemotherapy and hormone therapy publicly available because of concerns about completeness. However, chemotherapy and radiation therapy data can be made available to researchers by special request after signing a special data use agreement acknowledging an understanding of the data limitations.

Oncology Practice Claims and Linkage to Georgia SEER Data. The GCR obtained oncology claims from an oncology workflow solutions company that contracts with OP in almost every state in the United States to process health claims prior to submission to insurance companies for reimbursement. As such, these claims data include identifying information for the patient and provider, as well as date of service and diagnosis and procedure codes. The data were reported to the GCR under the legal framework of cancer surveillance reporting that exists in Georgia. Out of 12 OP in the state that utilize this single vendor for billing claim preparation, six agreed to report their cancer data to the GCR using a standardized billing claim format allowed by the registry. The GCR received retrospective claims data for the participating practices for 2013 to 2015 and prospective claims data via a live feed beginning in 2016.

SEER-Medicare Data. Individuals in the SEER data have been matched to the Medicare master enrollment file, which is maintained by the Centers for Medicare and Medicaid Services. Details of this linkage are described elsewhere (8). Briefly, the SEER-Medicare data include longitudinal enrollment and claims data (Part A: inpatient; Part B: physician and outpatient clinic) for Medicare beneficiaries who have been reported to a SEER registry as having had a cancer diagnosis. Providers are not required to submit Part A or B claims for beneficiaries who are enrolled in health maintenance organizations; therefore, inpatient and outpatient claims are available only for beneficiaries enrolled in fee-for-service plans. Similar to the OP claims, the Medicare claims also include information on date of service and diagnosis and procedure codes, including codes for chemotherapy administration.

Study Cohorts for Analyses. For this analysis, we constructed three study cohorts (see Figure 1). In cohort 1, we selected all individuals diagnosed with first primary cancer during 2013–2015 in the Georgia registry. Cohort 2 was developed by restricting cohort 1 patients to those having at least one OP claim within 12 months from their diagnosis date. To assess chemotherapy capture in OP, we then developed cohort 3, where we further restricted cohort 2 to patients aged 65 years and older enrolled in fee-for-service Medicare Part A and B from the diagnosis date through 12 months follow-up or to the date of death, whichever occurred first.

Statistical Analyses

The representativeness of the cancer patients included in the OP cohort in Georgia (cohort 2) was assessed by comparing those data against the demographic and clinical characteristics

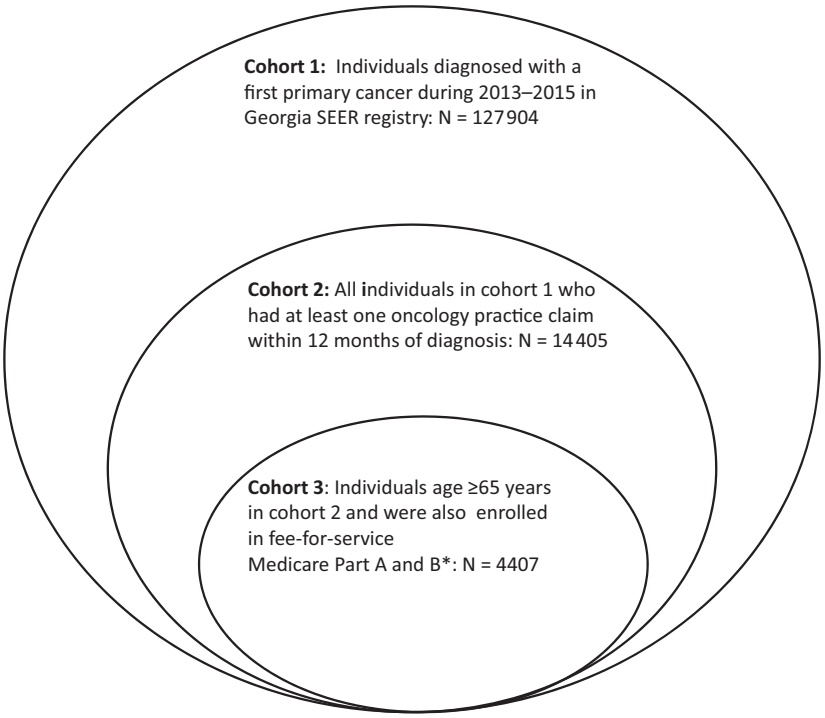

Figure 1. Construction of study cohorts for various analyses in terms of availability of data of different types. *From the date of diagnosis through 12 months follow-up or to date of death, whichever occurred first.

of all Georgia cancer patients in the catchment area (cohort 1). Cohort 3 consisted of patients who were found in both SEER-Medicare and the OP reporting to the GCR via their billing data. These data were used to assess the ability of the OP claims to capture chemotherapy in the patients under investigation, by comparing the data to Medicare claims (as gold standard) for chemotherapy. Flag variables were created for each patient to indicate whether they had received chemotherapy (yes or no) per the OP and Medicare data. We used a master list of codes compiled by the NCI SEER Program to define chemotherapy receipt (codes list found at https://seer.cancer.gov/oncologytoolbox/) as of January 1, 2019. In defining the chemotherapy flag, we attempted to follow the existing rules used by registries to capture first course of therapy. For example, we did not count radiosensitizing chemotherapy (n = 11) toward the algorithm, because registries typically do not count this as part of first course of chemotherapy. With Medicare data serving as the gold standard, sensitivity, specificity, and kappa statistics for chemotherapy receipt per OP claims were calculated by demographic and clinical characteristics. Statistics were also calculated using varying time windows for assessing receipt of chemotherapy to determine whether agreement between Medicare and OP data was time dependent. This was done mainly because SEER does not define a fixed interval for the collection of initial chemotherapy course. So, we wanted to assess the time window (eg, 6 months from diagnosis) within which most of the chemotherapy is administered following a cancer diagnosis. To evaluate this, the overall sensitivity, specificity, and kappa statistics were calculated using a 3-, 6-, and 9-month postdiagnosis Medicare and OP claims window to provide a comparison to the 12-month claims window.

The study was approved by the Institutional Review Board at Georgia Department of Public Health.

Results

Our analysis included a total of 127 904 individuals diagnosed with an invasive, first primary cancer in the GCR over the 2013–2015 period (cohort 1). Of this total, 14 405 (11.2%) individuals had at least one OP claim within 12 months of their cancer diagnosis (cohort 2), and 4460 (3.4%) individuals had at least one OP claim for chemotherapy during this time period (Table 1). In comparison to all cancer patients included in the Georgia registry (ie, compare column C to column A, Table 1), characteristics differed among those included in the OP data as follows: white (75.8% vs 68.0%), female (59.7% vs 50.5%), insured (82.1% vs 72.6%), high SES (census tract third quintile income: 28.7% vs 24.9%), and stage IV disease (23.9% vs 17.4%). In a subgroup analysis, we again compared these demographic characteristics among a subset of patients identified as receiving chemotherapy (as per the registry-documented chemotherapy variable) to the subset of patients receiving chemotherapy according to OP claims (ie, compare column D to column B, Table 1). Results show that even when we restrict to only those cancer patients receiving chemotherapy, the practices still underrepresented non-Hispanic blacks (19.3% vs 28.9%;) and other minorities, uninsured (1.1% vs 6.7%) and slightly overrepresented females (60.1% vs 55.9%), patients with third-quintile income levels (29.5% vs 25.2%), and patients residing outside the metropolitan Atlanta area (79.5% vs 67.1%).

A total of 4407 Georgia patients were diagnosed in 2013–2015 who had an OP claim within 12 months of diagnosis and also had relevant Medicare data during this time period (cohort 3). Overall, in comparison to Medicare claims, the sensitivity of OP data to identify individuals who received chemotherapy within 12 months of diagnosis was 76.1% (Table 2). There was no indication that the ability of the OP data to capture chemotherapy was dependent on the time period assessed. In fact, we found

Table 1: Characteristics of Georgia SEER cancer cases diagnosed between 2013-2015 overall, among those identified having chemotherapy by registry documented chemotherapy variable, among those who linked to Oncology Practice, and among those who were identified as receiving chemotherapy according to the Oncology Practice Claims.

	Column A	Column B	Column C	Column D
	All GA Cases	Received chemotherapy according to registry documented chemotherapy variable	Among GA cases with any claims from oncology practice[1]	Received chemotherapy according to oncology practice[2]
	No. (%)	No. (%)	No. (%)	No. (%)
Total	127904 (100)	40479 (100)	14405 (100)	4460 (100)
Year of Diagnosis				
2013	41235 (32.24)	13150 (32.49)	4586 (31.84)	1533 (34.37)
2014	42614 (33.32)	13557 (33.49)	4800 (33.32)	1508 (33.81)
2015	44055 (33.44)	13772 (34.02)	5019 (34.84)	1419 (31.82)
Age at diagnosis, years				
<45	13686 (10.70)	5754 (14.21)	1137 (7.89)	446 (10.00)
45-54	20157 (15.76)	7785 (19.23)	2162 (15.01)	757 (16.97)
55-64	34565 (27.02)	11897 (29.39)	3736 (25.94)	1221 (27.38)
65-74	35387 (27.67)	10446 (25.81)	4442 (30.84)	1444 (32.38)
75+	24109 (18.85)	4597 (11.36)	2928 (20.33)	592 (13.27)
Race/Ethnicity				
NH White	87081 (68.08)	26135 (64.56)	10924 (75.83)	3460 (77.58)
NH Black	33693 (26.34)	11706 (28.92)	2995 (20.79)	864 (19.37)
NH AI/AN	114 (0.09)	47 (0.12)	11 (0.08)	2 (0.04)
NH API	2411 (1.89)	954 (2.36)	152 (1.06)	41 (0.92)
Hispanic	4029 (3.15)	1612 (3.98)	299 (2.08)	90 (2.02)
Unknown	576 (0.45)	25 (0.06)	24 (0.17)	3 (0.07)
Sex				
Male	63213 (49.42)	17850 (44.1)	5793 (40.22)	1776 (39.82)
Female	64691 (50.58)	22629 (55.9)	8612 (59.78)	2684 (60.18)
Insurance Status[3]				
Uninsured	6200 (4.85)	2743 (6.78)	328 (2.28)	51 (1.14)
Medicaid	12699 (9.93)	5410 (13.36)	1401 (9.73)	373 (8.36)
Insured	92915 (72.64)	31485 (77.78)	11829 (82.12)	3899 (87.42)
Unknown	16090 (12.58)	841 (2.08)	847 (5.88)	137 (3.07)
Region				
Metropolitan Atlanta	39874 (31.17)	12661 (31.28)	2266 (15.73)	797 (17.87)
Greater Georgia	85925 (67.18)	27177 (67.14)	11725 (81.40)	3547 (79.53)
Rural Georgia	2105 (1.65)	641 (1.58)	414 (2.87)	116 (2.60)
Census Tract level income				
Q1: $5968-$37,655	31996 (25.02)	10641 (26.29)	3120 (21.66)	847 (18.99)
Q2: $37,813-$50,549	31971 (25.00)	10352 (25.57)	3897 (27.05)	1229 (27.56)
Q3: $50,572-$68,235	31962 (24.99)	10207 (25.22)	4147 (28.79)	1318 (29.55)
Q4: $68,341-$1,87,750	31955 (24.98)	9275 (22.91)	3240 (22.49)	1066 (23.90)
Unknown	20 (0.02)			
Stage at diagnosis[4]				
I	35416 (27.69)	4578 (11.31)	3275 (22.74)	563 (12.62)
II	23787 (18.60)	7824 (19.33)	2594 (18.01)	1008 (22.60)
III	15194 (11.88)	9466 (23.38)	2345 (16.28)	1143 (25.63)
IV	22305 (17.44)	12293 (30.37)	3454 (23.98)	1295 (29.04)
Unknown	18683 (14.61)	5891 (14.55)	1931 (13.41)	442 (9.91)
Cancer Site[5]				
Breast	21681 (16.95)	8232 (20.34)	4047 (28.09)	1262 (28.30)
Lung	14938 (11.68)	6823 (16.86)	2308 (16.02)	880 (19.73)
Colon	7542 (5.90)	2797 (6.91)	964 (6.69)	337 (7.56)
NHL	4352 (3.40)	2496 (6.17)	698 (4.85)	284 (6.37)
Uterus	3622 (2.83)	984 (2.43)	310 (2.15)	66 (1.48)
Rectum	3352 (2.62)	1833 (4.53)	476 (3.30)	196 (4.39)
Oral	3292 (2.57)	1628 (4.02)	404 (2.80)	165 (3.70)
Pancreas	3240 (2.53)	1639 (4.05)	459 (3.19)	200 (4.48)
Myeloma	2129 (1.66)	1237 (3.06)	307 (2.13)	148 (3.32)
Other	63756 (49.85)	12810 (31.65)	4432 (30.77)	922 (20.67)

AI: American Indian; AN: Alaska Native; NH: Non-Hispanic; NHL: Non-Hodgkin Lymphoma; SEER: Surveillance, Epidemiology and End Results

1Cancer cases who were in Georgia registry who had at least one claim of any types in the private oncology practice within the first 12 months of cancer diagnosis.

2Cancer cases who were in Georgia registry who had at least one chemotherapy claim in the private oncology practice within the first 12 months of cancer diagnosis.

3Median household income level of residential census tract at the time of cancer diagnosis.

4American Joint Committee on Cancer, 7th edition.

5For cancer sites with less than 100 cases were collapsed into “other cancer” group.

For unknown census tract level income the numbers are not shown due to less than 16 cases in the cell.

Table 2. Accuracy of oncology practice (OP) claims to identify receipt of chemotherapy overall and by time window since diagnosis among cancer patients diagnosed in 2013–2015 as reported to the Georgia SEER registry, in comparison to gold-standard Medicare claims*

Time window of diagnosis for capturing chemotherapy	Total	Medicare, Yes OP, Yes	Medicare, Yes OP, No	Medicare, No OP, Yes	Medicare, No OP, No	Sensitivity,%	Specificity, %	Kappa
0–3 month window	4407	1003	317	28	3059	76.0%	99.1%	0.80
0–6 month window	4407	1104	343	31	2929	76.3	99.0	0.80
0–9 month window	4407	1138	353	31	2885	76.3	98.9	0.79
0–12 month window	4407	1150	361	31	2865	76.1	98.9	0.79

*We restricted cases to SEER sequence number = 0 or 1 (ie, restricted to those diagnosed with only one or first tumor). SEER = Surveillance, Epidemiology and End Results.

that most of the chemotherapy was given within 4 months of diagnosis because increasing the time window to 6-, 9- or 12-month postdiagnosis changed the results very little (Table 2).

The sensitivity of the OP data to identify chemotherapy varied by demographic and clinical characteristics (Table 3). For example, starting at ages 65–69 years, sensitivity increased until ages 80–84 years (74.6%, 75.3%, 77.9%, and 79.4%), but then decreased slightly among patients aged 85 years and older (78.3%; Table 3). Sensitivity was also higher among non-Hispanic white patients than non-Hispanic blacks (78.3% vs 63.2%; Table 3). In terms of clinical characteristics, sensitivity varied considerably by cancer site (oral cancer: 83.0%; uterus cancer: 46.2%; Table 3) and by stage (stage I: 72.0%; stage IV: 73.7%; Table 3). There was very little variation in sensitivity by sex and region.

There was very little gained by adding the single chemotherapy (yes or no) information from OP claims to the chemotherapy information already collected by the registry. For example, documented chemotherapy use went up from 29.4% to 29.7% after updating the SEER chemotherapy variable using the OP claims data. This was an encouraging finding in that these data are already being captured from other reporting sources across the state. Additionally, we have found that among the 4460 patients identified from OP claims as receiving chemotherapy, 4184 patients (94%) also had documented chemotherapy by the registry chemotherapy variable.

Discussion

Increasingly, cancer patients are receiving care across a wide range of clinical settings and specialty practices. As a result, studies of real-world cancer treatment for patients, outside the clinical trial setting, have been constrained by the fragmentation of data across health-care systems. With greater availability of electronic health records, cancer surveillance has the opportunity to assess the potential of other data resources to close information gaps. Claims are an attractive option for linkage with cancer registries, because they are becoming more available electronically, can be transferred in almost real time, and have structured information. However, assessments of the usefulness of such data to enhance cancer registry data have been limited to studies using SEER-Medicare data. We found that cancer patients who sought cancer care in the six Georgia private OP included in our analysis were not representative of the underlying patient population in the GCR. Furthermore, we found that the practices underrepresented minorities and uninsured cancer patients in Georgia and overrepresented females, persons with high SES, and patients residing outside the metropolitan Atlanta area. Some of these differences between the patients in our study and the general cancer population are expected. For example, patients for whom chemotherapy is not standardly recommended because of the cancer type or an early stage are underrepresented in the OP data, because some of these patients did not require care from an oncologist. Another large group of patients, men with prostate cancer, are largely treated by urologists and would not be expected to appear in OP data. Additionally, the ability of the OP data to identify chemotherapy was moderate (76.1% sensitivity) in GCR but varied somewhat by demographic and clinical characteristics (76.1% to 83.0%, respectively).

As such, attempting to use these OP data alone to assess treatment will miss a number of patients who have been treated with chemotherapy in other locations but still appear in the billing claims because of OP visits for other purposes. This will result in misclassification and inaccurate estimates of the proportion of individuals treated among cancer patients in the Georgia SEER catchment area. We attempted to investigate why patients had a claim for chemotherapy in the SEER-Medicare data but had no claim in the OP data. We found that in the Medicare data, the majority of these patients were found to have received chemotherapy in the hospital outpatient settings and not in the physicians' offices, thus their absence from the OP claims. The acquisition of OP by hospitals is increasing because of changes in payment policies.

There are a few strengths and limitations of our study. One of the challenges is the lack of representativeness, because our study has shown that cancer patients who seek cancer care in these six OPs are not representative of the registry cancer patient population. These findings demonstrate that one cannot make population-based inference based on these data. Obtaining consent and claims from a larger number of practices has the potential to make the data somewhat more representative, but these efforts would be limited if they did not include practices that serve a more diverse group of patients. Also, statistical methods (eg, propensity scores) could be used to adjust the result to be representative of the registry cancer population (12, 13). The second and more important challenge is that some patients identified in the OP claims actually received their chemotherapy elsewhere; as a result, patients who received chemotherapy that was not reported in the OP data cannot be differentiated from patients who did not receive any chemotherapy. This limits the potential of these data to be used in patient-level studies of the patterns of cancer care, that is, analyses characterizing chemotherapy use among different patient subgroups. Nonetheless, the OP data have many strengths and could be used for other purposes. For example, because claims are received in almost real time, they are currently being used to update follow-up information in the registry. Additionally, these claims could be used for rapid case ascertainment and to flag cancer patients for potential recurrence, provided that patients have been diagnosed years ago and are found to have sought cancer care recently. Furthermore, these claims can be

Table 3. Accuracy of oncology practice (OP) claims to identify receipt of chemotherapy overall and by demographic and clinical characteristics among cancer patients diagnosed in 2013–2015 as reported to the Georgia SEER registry, in comparison to gold-standard Medicare claims*

Characteristics	Total	Sensitivity, %	Specificity, %	Kappa
Age at diagnosis, y				
65–69	1293	74.6	97.8	0.74
70–74	1118	75.3	98.5	0.77
75–79	915	77.9	99.5	0.82
80–84	642	79.4	99.6	0.84
≥85	439	78.3	100.0	0.87
Race/ethnicity				
Non-Hispanic white	3750	78.3	99.0	0.81
Non-Hispanic black	553	63.2	98.3	0.66
Other	104	69.0	98.7	0.74
Sex				
Female	2466	74.9	99.3	0.79
Male	1941	77.5	98.5	0.80
Region				
Greater Georgia	3598	74.6	98.9	0.78
Metropolitan Atlanta	667	86.2	98.7	0.88
Rural Georgia	142	75.0	100.0	0.81
Census tract level income				
Q1: $5968–$37 655	847	69.2	98.4	0.73
Q2: $37 813–$50 549	1293	74.5	98.7	0.77
Q3: $50 572–$68 235	1263	80.9	99.4	0.84
Q4: $68 341–$187 750	1004	78.0	99.1	0.81
Stage at diagnosis†				
I	968	72.0	100.0	0.81
II	745	80.9	99.2	0.83
III	684	77.3	96.1	0.71
IV	1141	73.7	98.1	0.74
Unstaged	692	76.5	99.2	0.81
Cancer site‡				
Breast	920	82.0	99.7	0.87
Colon	320	77.0	98.1	0.79
Rectum	116	69.6	93.3	0.63
Lung	893	73.4	97.9	0.74
Myeloma	95	80.4	93.2	0.73
Non-Hodgkin lymphoma	237	77.2	100.0	0.78
Oral cavity and pharynx	94	83.0	100.0	0.81
Pancreas	170	83.0	96.3	0.79
Uterus	96	46.2	100.0	0.44
Other cancers	1466	76.8	99.6	0.82

*We restricted cases to SEER sequence number = 0 or 1 (ie, restricted to those diagnosed with only one or first tumor). SEER = Surveillance, Epidemiology and End Results.

†*American Joint Committee on Cancer*, 7th edition.

‡For cancer sites with less than 100, cases were collapsed into "other cancer" group.

used for identification and reporting of secondary or longitudinal treatment beyond the initial course of therapy (yes or no) and can also provide information on agents, doses, and receipt dates. These claims contain information on cancer patients younger than 65 years, which is not available in SEER-Medicare.

Restricting the data to only cases that have chemotherapy claims could be used for generating hypotheses, such as those related to quality of care (14). For example, treatment delays, patterns of drug use by race and ethnicity and socioeconomic status, use of ancillary medications, or novel diagnostics to target chemotherapy (eg, tumor genomic profiling for recurrence risk) could be addressed with such a data source. Inference should be limited to this cohort of patients. Nonetheless, these questions are key priorities for future research.

In conclusion, we used an existing stream of health claims to identify patients in the Georgia SEER registry and assess the utility of the health claims to capture chemotherapy use. Although we found underreporting of chemotherapy use from these OP, the data can be valuable for enhancing cancer registry information in many other ways, such as updating the follow-up information on registry cases and rapid case findings. However, to fulfill the potential as an independent cancer research database resource, the representativeness of the OP claims data needs to be improved by including other registries, and their completeness needs to be improved by including hospital-based practices or other sources for chemotherapy information.

Funding

The collection of cancer incidence data in Georgia was supported by contract HHSN261201800003I, Task Order HHSN26100001 from the NCI and cooperative agreement 5 NU58DP006352-03–00 from the CDC.

Notes

Affiliations of authors: Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD (NH, ABM); Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA (KCW); Special Volunteer, Healthcare Delivery Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD (JLW); Information Management Services, Rockville, MA (DSC, LC).

Disclaimer: Findings and conclusions are the authors' and do not necessarily represent the official positions of their affiliations or those of the National Cancer Institute, the National Institutes of Health, or the US Department of Health and Human Services.

References

1. Howlader NA, Krapcho M, Miller D, et al. SEER Cancer Statistics Review, 1975-2016, National Cancer Institute, based on November 2018 SEER data submission, posted to the SEER web site. April 2019. https://seer.cancer.gov/csr/1975_2016/. Accessed June 15, 2019.
2. Noone AM, Lund JL, Mariotto A, et al. Comparison of SEER treatment data with Medicare claims. *Med Care*. 2016;54(9):e55–64.
3. Do YK, Carpenter WR, Spain P, et al. Race, healthcare access and physician trust among prostate cancer patients. *Cancer Causes Control*. 2010;21(1):31–40.
4. O'Malley AS, Forrest CB, Feng S, et al. Disparities despite coverage: gaps in colorectal cancer screening among Medicare beneficiaries. *Arch Intern Med*. 2005;165(18):2129–2135.
5. Hoffman C. The role of health insurance coverage in reducing racial/ethnic disparities in health care. *Health Aff*. 2005;24(2):398–408.
6. Souza JAD Jr. Medicare reimbursement changes and the practice of oncology: understanding of the past is a key to the future. *J Oncol Pract*. 2011;7(5):306–308.
7. Lund JL, Stürmer T, Harlan LC, et al. Identifying specific chemotherapeutic agents in Medicare data: a validation study. *Med Care*. 2013;51(5):e27–e34.
8. Warren JL, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care*. 2002;40(suppl):IV–IV18.
9. Lamont EB, Herndon JE, Weeks JC, et al. Measuring clinically significant chemotherapy-related toxicities using Medicare claims from cancer and leukemia group B (CALGB) trial participants. *Med Care*. 2008;46(3):303–308.
10. Jagsi R, Abrahamse P, Hawley ST, et al. Underascertainment of radiotherapy receipt in surveillance, epidemiology, and end results registry data. *Cancer*. 2012;118(2):333–341.
11. Cooper GS, Yuan Z, Stange KC, et al. Agreement of Medicare claims and tumor registry data for assessment of cancer-related treatment. *Med Care*. 2000;38(4):411–421.
12. Stuart EA, Bradshaw CP, Leaf P. Assessing the generalizability of randomized trial results to target populations. Prev Sci. 2015;16(3):475–485.
13. Stuart EA, Cole SR, Bradshaw CP, et al. The use of propensity scores to assess the generalizability of results from randomized trials. *J R Stat Soc Ser A*. 2011;174(2):369–386.
14. Kurian AW, Lichtensztajn DY, Keegan THM, et al. Patterns and predictors of breast cancer chemotherapy use in Kaiser Permanente Northern California, 2004–2007. Breast Cancer Res Treat. 2013;137(1):247–260.