The Oncologist®

Data Linkage to Improve Geriatric Oncology Research: A Feasibility Study

Jennifer L. Lund,[a,b] Anne-Marie Meyer,[a,b] Allison M. Deal,[b] Bong-Jin Choi,[b] YunKyung Chang,[b] Grant R. Williams,[b,c,d] Mackenzi Pergolotti,[b,c,e] Emily J. Guerard,[c,f] Hyman B. Muss,[b,c] Hanna K. Sanoff[b,c]

[a]Department of Epidemiology, Gillings School of Global Public Health, [b]Lineberger Comprehensive Cancer Center and [c]Division of Hematology/Oncology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; [d]Division of Hematology/Oncology, School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA; [e]Department of Occupational Therapy, College of Health and Human Sciences, Colorado State University, Fort Collins, Colorado, USA; [f]Division of Hematology/Oncology, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin, USA
Disclosures of potential conflicts of interest may be found at the end of this article.

ABSTRACT

Older adults (aged 65 years and older) diagnosed with cancer account for most cancer-related morbidity and mortality in the United States but are often underrepresented on clinical trials. Recent attention from a variety of professional, research, regulatory, and patient advocacy groups has centered on data linkage and data sharing as a means to capture patient information and outcomes outside of clinical trials to accelerate progress in the fight against cancer. The development of a more robust observational research data infrastructure would help to address gaps in the evidence base regarding optimal approaches to treating cancer among the growing and complex population of older adults. To demonstrate the feasibility of building such a resource, we linked information from a sample of older adults with cancer in North Carolina using three distinct, but complementary, data sources: (a) the Carolina Senior Registry, (b) the North Carolina Central Cancer Registry, and (c) North Carolina fee-for-service Medicare claims data. A description of the linkage process, metrics, and characteristics of the final cohort is reported. This study highlights the potential for data linkage to improve the characterization of health status among older adults with cancer and the possibility to conduct passive follow-up for outcomes of interest over time. Extensions of these linkage efforts in partnership with other institutions will enhance our ability to generate evidence that can inform the management of older adults with cancer. *The Oncologist* 2017;22:1002–1005

INTRODUCTION

Progress in the fight against cancer has relied heavily on evidence generated from clinical trials, representing the experience of less than 5% of all newly diagnosed cancer patients [1]. Older adults (age 65+ years) are disproportionately affected by this trial-driven research paradigm, as they are less likely than younger patients to be enrolled onto clinical trials [1, 2] despite accounting for the majority of cancer-related morbidity and mortality in the United States [3].

National efforts led by the American Society of Clinical Oncology (ASCO), the National Cancer Institute, the Food and Drug Administration, and most recently by former Vice President Joe Biden and the Cancer Moonshot initiative, have promoted data linkage and data sharing as a means to capture patient information and outcomes outside of clinical trials to accelerate progress in the fight against cancer. In a recent position statement on improving the evidence base for treating older adults [4], ASCO recommended developing a robust research infrastructure outside of clinical trials using observational methods and collection of relevant geriatric oncology-specific information alongside other patient, tumor, and treatment data.
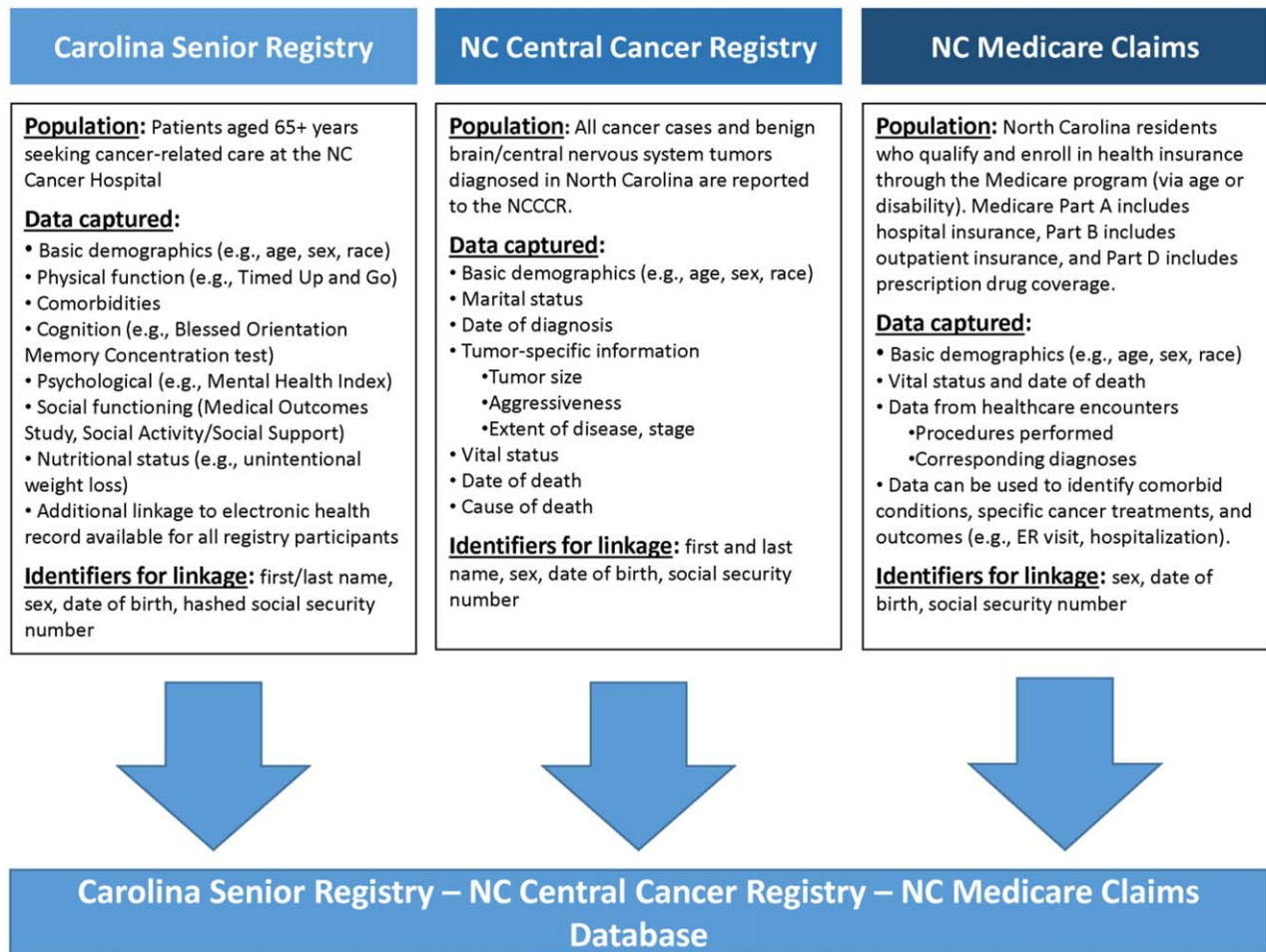
To demonstrate the feasibility of building such a resource for geriatric oncology, we linked information from a sample of older adults with cancer in North Carolina using three distinct data sources.

MATERIALS AND METHODS

The Carolina Senior Registry (NCT01137825) is a hospital-based cancer registry including one academic and seven community sites, started in 2009 to collect data from a brief geriatric assessment performed on older adults living with cancer [5], modeled after the tool developed by Hurria and colleagues [6]. Geriatric assessment data were collected through healthcare provider evaluation (performed by the treating oncologist or trained clinical research assistant) and patient-administered questionnaires. The current study is limited to patients treated at the one academic site, accounting for 69% of all patients in

Carolina Senior Registry

Population: Patients aged 65+ years seeking cancer-related care at the NC Cancer Hospital

Data captured:
• Basic demographics (e.g., age, sex, race)
• Physical function (e.g., Timed Up and Go)
• Comorbidities
• Cognition (e.g., Blessed Orientation Memory Concentration test)
• Psychological (e.g., Mental Health Index)
• Social functioning (Medical Outcomes Study, Social Activity/Social Support)
• Nutritional status (e.g., unintentional weight loss)
• Additional linkage to electronic health record available for all registry participants

Identifiers for linkage: first/last name, sex, date of birth, hashed social security number

NC Central Cancer Registry

Population: All cancer cases and benign brain/central nervous system tumors diagnosed in North Carolina are reported to the NCCCR.

Data captured:
• Basic demographics (e.g., age, sex, race)
• Marital status
• Date of diagnosis
• Tumor-specific information
  • Tumor size
  • Aggressiveness
  • Extent of disease, stage
• Vital status
• Date of death
• Cause of death

Identifiers for linkage: first and last name, sex, date of birth, social security number

NC Medicare Claims

Population: North Carolina residents who qualify and enroll in health insurance through the Medicare program (via age or disability). Medicare Part A includes hospital insurance, Part B includes outpatient insurance, and Part D includes prescription drug coverage.

Data captured:
• Basic demographics (e.g., age, sex, race)
• Vital status and date of death
• Data from healthcare encounters
  • Procedures performed
  • Corresponding diagnoses
• Data can be used to identify comorbid conditions, specific cancer treatments, and outcomes (e.g., ER visit, hospitalization).

Identifiers for linkage: sex, date of birth, social security number

Carolina Senior Registry – NC Central Cancer Registry – NC Medicare Claims Database

Figure 1. Patient information contained in each of the three data sources used in the linkage feasibility study of older adults with cancer in North Carolina.
Abbreviations: ER, emergency room; NC, North Carolina; NCCCR, North Carolina Central Cancer Registry.

the Carolina Senior Registry, as all required data elements for the linkage were readily available.

Using patients' first and last name, date of birth, sex, and hashed social security number, we applied deterministic and probabilistic linkage algorithms to patients in the Carolina Senior Registry and the North Carolina Central Cancer Registry (NCCCR), which captures tumor information legally reportable by all healthcare providers in North Carolina. We then used a crosswalk developed by UNC Lineberger Integrated Cancer Information and Surveillance System researchers to map patients in the NCCCR to the Medicare enrollment database [7]. Medicare fee-for-service claims capture longitudinal information about beneficiaries' healthcare encounters, including diagnoses and procedures from hospitalizations and outpatient visits.

Figure 1 provides an overview of the information contributed by each data source. Analyses were conducted using SAS version 9.4 (Cary, North Carolina). Descriptive statistics for the cohort are presented. This study was approved by the Institutional Review Board at the University of North Carolina at Chapel Hill.

RESULTS

Of the 703 older cancer patients in the Carolina Senior Registry, we linked 636 patients (90%) to the NCCCR and 546 (78%) to

the Medicare enrollment database. Among this population, 369 (68%) were continuously enrolled in Medicare Parts A and B (without managed care enrollment) for 6 months prior to the geriatric assessment date. Characteristics of the study population stratified by claims-based Charlson comorbidity score [8] are reported in Table 1.

The study population primarily includes women, driven largely by overrepresentation of early-stage breast cancer patients seen by recruiting oncologists. Median age was 74 years, and more than half of the population had some college education. The prevalence of patients with unintentional weight loss, prolonged Timed-Up and Go (less than14 seconds, indicating poor mobility), or a physician-rated Karnofsky Performance Status of less than 80% steadily increased with a higher Medicare claims-based Charlson comorbidity score. However, the prevalence of patients with polypharmacy (9+ medications) or a Blessed Orientation Memory Concentration test score of less than 11 (suggestive of memory impairment) did not monotonically increase with higher Charlson comorbidity score. These preliminary results demonstrate how linked data resources enrich our ability to identify important variation in health status among older adults with cancer that cannot be described using a singular measure.

Table 1. Characteristics of sample of older adults, overall and by Charlson Comorbidity Index

Characteristics	Overall	Charlson Comorbidity Index (CCI) calculated from Medicare claims		
		CCI = 0	CCI = 1	CCI = 2+
Carolina Senior Registry Data, n	369	251	70	48
Time of geriatric assessment				
Within 90 days from cancer diagnosis, n (%)	133 (36)	88 (35)	26 (37)	19 (40)
90–365 days from cancer diagnosis, n (%)	75 (20)	46 (18)	15 (21)	14 (29)
>365 days from cancer diagnosis, n (%)	161 (44)	117 (47)	29 (41)	15 (31)
Age at geriatric assessment, mean (range)	74 (65–100)	74 (65–100)	74 (66–94)	75 (66–93)
65–70 years, n (%)	132 (36)	96 (38)	23 (33)	13 (27)
70–74 years, n (%)	110 (30)	73 (29)	22 (31)	15 (31)
75+ years, n (%)	127 (34)	82 (33)	25 (36)	20 (42)
Sex				
Male, n (%)	90 (24)	53 (21)	21 (30)	16 (33)
Female, n (%)	279 (76)	198 (79)	49 (70)	32 (67)
Race				
White, n (%)	323 (88)	–	–	–
Non-White, n (%)	45 (12)	–	–	–
Highest educational level				
High school degree or less*, n (%)	174 (47)	107 (43)	39 (56)	28 (58)
Some college education, n (%)	195 (53)	144 (57)	31 (44)	20 (42)
Body mass index, median (IQR)	26 (23, 30)	26 (23, 30)	26 (23, 30)	28 (25, 32)
Selected geriatric assessment measures				
Physical function score, median (IQR)	15 (10, 18)	16 (12, 19)	11 (7, 16)	8 (5, 16)
BOMC ≥11, median (IQR)	4 (2, 6)	2 (2, 6)	2 (2, 6)	6 (2, 8)
MHI 13 ≥12 or anxious (depressed), n (%)	60 (16)	–	–	–
MHI 13 ≥6 (anxious), n (%)	88 (24)	56 (22)	18 (26)	14 (29)
Daily medications (≥9 medications), n (%)	143 (39)	83 (33)	36 (51)	24 (50)
Reported falls in past 6 months (1+ fall), n (%)	88 (24)	50 (20)	16 (23)	16 (33)
Unintentional weight loss ≥5%, n (%)	80 (22)	41 (16)	20 (29)	19 (40)
TUG ≥14 seconds (or unable), n (%)	113 (31)	64 (25)	26 (37)	23 (48)
Physician-rated KPS ≤80%, n (%)	43 (12)	16 (6)	14 (20)	13 (27)
NCCCR Data				
Cancer site				
Breast, n (%)	194 (53)	152 (61)	29 (41)	13 (27)
Hematologic malignancies, n (%)	47 (13)	25 (10)	11 (16)	11 (23)
Other, n (%)	128 (35)	74 (29)	30 (43)	24 (50)
AJCC Stage 6th Edition, n (%)				
0–II, n (%)	198 (54)	148 (59)	29 (41)	21 (44)
III–IV, n (%)	120 (33)	74 (29)	30 (43)	16 (33)
Missing or unstaged, n (%)	51 (14)	29 (12)	11 (16)	11 (23)
Medicare Enrollment and Claims Data				
Deaths, n (%)	92 (25)	53 (21)	24 (34)	15 (31)
Number of months from GA to death or 12/31/2013, median (IQR)	19 (7, 34)	23 (11, 38)	13 (5, 27)	12 (5, 23)

[a]Indicates that missing values were added to this category.

Abbreviations: –, at least one cell size was less than 11 and therefore had to be suppressed for confidentiality; AJCC, American Joint Commission on Cancer; BOMC, Blessed Orientation Memory Concentration test; CCI, Charlson Comorbidity Index; KPS, Karnofsky Performance Score; GA, geriatric assessment; IQR, interquartile range; MHI, Mental Health Index; TUG, Timed-Up and Go.

DISCUSSION

We demonstrated the feasibility of linking information from three disparate but complimentary data sources for a sample of older adults with cancer treated at a North Carolina academic medical center. This data resource richly characterized the health status and tumor features of older adults living with cancer and established a mechanism to identify specific treatments administered and conduct passive surveillance and longitudinal

follow-up using Medicare enrollment and claims data. As cancer treatment regimens become increasingly complex, both regulators and patient advocacy groups want to better understand the impact of treatment on patient-centered geriatric oncology outcomes (e.g., emergency department visits, nursing home placement, decline in physical function, memory impairment), which are not traditionally captured in clinical trials. Therefore, linkage of rich but disparate data sources represents a promising avenue to capture more multifaceted outcome measures.

Other groups have launched research initiatives to improve the evidence base for treating older adults with cancer or collect real-world data on cancer patients outside of the trial setting. For example, the Cancer and Aging Research Group has undertaken numerous studies collecting geriatric assessment data on patients from multiple medical centers [9]. However, these data have not yet been linked to state cancer registries or Medicare, Medicaid, or other insurance claims data, limiting their ability to passively follow patients over time. On the other hand, researchers within the Cancer Research Network have developed a robust infrastructure for pooling electronic health record data for cancer patients across multiple integrated healthcare delivery systems [10], but detailed geriatric assessment data are not routinely collected for older patients with cancer in these settings.

Conclusion

While the research questions that can be explored using the linked Carolina Senior Registry data are limited by the small sample and limited generalizability at present, we plan to partner with institutions interested in conducting similar linkages in order to build a larger database well-suited to address a variety of questions regarding the care of older adults with cancer. In the future, large, linked data sources can be used to fill existing evidence gaps regarding long-term benefits and harms of cancer treatments, and ultimately to tailor the delivery of cancer care and improve outcomes for older adults with cancer.

Author Contributions

Conception/Design: Jennifer L. Lund, Anne-Marie Meyer, Mackenzi Pergolotti, Hyman B. Muss, Hanna K. Sanoff
Provision of study material or patients: Jennifer L. Lund, Anne-Marie Meyer, Bong-Jin Choi, Hyman B. Muss, Hanna K. Sanoff
Collection and/or assembly of data: Jennifer L. Lund, Anne-Marie Meyer, Allison M. Deal, Bong-Jin Choi, YunKyung Chang
Data analysis and interpretation: Jennifer L. Lund, Anne-Marie Meyer, Allison M. Deal, YunKyung Chang, Grant R. Williams, Mackenzi Pergolotti, Emily J. Guerard, Hyman B. Muss, Hanna K. Sanoff
Manuscript writing: Jennifer L. Lund, Allison M. Deal, Grant R. Williams, Emily J. Guerard, Hanna K. Sanoff

Disclosures

Jennifer L. Lund: GlaxoSmithKline (E [spouse]); **Anne-Marie Meyer**: Merck (H); **Hanna K. Sanoff**: Bayer, Novartis, Precision Biologics, Immunomedics, Merck (RF). The other authors indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

References

1. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: Race-, sex-, and age-based disparities. JAMA 2004;291:2720–2726.

2. Talarico L, Chen G, Pazdur R. Enrollment of elderly patients in clinical trials for cancer drug registration: A 7-year experience by the US Food and Drug Administration. J Clin Oncol 2004;22:4626–4631.

3. Howlader N, Noone AM, Krapcho M et al. SEER Cancer Statistics Review, 1975—2012, Available at http://seer.cancer.gov/csr/1975_2012/. Accessed October 1, 2016.

4. Hurria A, Levit LA, Dale W et al. Improving the Evidence Base for Treating Older Adults With Cancer: American Society of Clinical Oncology Statement. J Clin Oncol 2015;33:2826–3833.

5. Williams GR, Deal AM, Jolly TA et al. Feasibility of geriatric assessment in community oncology clinics. Geriatr Oncol 2014;5:245–251.

6. Hurria A, Gupta S, Zauderer M et al. Developing a cancer-specific geriatric assessment: A feasibility study. Cancer 2005;104:1998–2005.

7. Meyer AM, Olshan AF, Green L et al. Big data for population-based cancer research: The integrated cancer information and surveillance system. N C Med J 2014;75:265–269.

8. Klabunde CN, Legler JM, Warren JL et al. A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. Ann Epidemiol 2007;177: 584–590.

9. Hurria A, Lichtman SM. Clinical pharmacology of cancer therapies in older adults. Br J Cancer 2008; 98:517–522.

10. Chubak J, Ziebell R, Greenlee RT et al. The Cancer Research Network: A platform for epidemiologic and health services research on cancer prevention, care, and outcomes in large, stable populations. Cancer Causes Control 2016;27:1315–1323.