Published in final edited form as:

JCO Clin Cancer Inform. 2019 May ; 3: 1–8. doi:10.1200/CCI.19.00011.

Identifying Smoking Status and Smoking Cessation Using a Data Linkage Between the Kentucky Cancer Registry and Health Claims Data

Michael Shayne Gallaway, PhD, MPH[1], **Bin Huang, DrPH**[2], **Quan Chen, DrPH**[2], **Tom Tucker, PhD, MPH**[2], **Jaclyn McDowell, DrPH**[2], **Eric Durbin, DrPH**[2], **David Siegel, MD**[1], **Eric Tai, MD**[1]

[1]Centers for Disease Control and Prevention, Atlanta, GA

[2]University of Kentucky, Lexington, KY

Abstract

PURPOSE—Linkage of cancer registry data with complementary data sources can be an informative way to expand what is known about patients and their treatment and improve delivery of care. The purpose of this study was to explore whether patient smoking status and smoking-cessation modalities data in the Kentucky Cancer Registry (KCR) could be augmented by linkage with health claims data.

METHODS—The KCR conducted a data linkage with health claims data from Medicare, Medicaid, state employee insurance, Humana, and Anthem. Smoking status was defined as documentation of personal history of tobacco use (International Classification of Diseases, Ninth Revision [ICD-9] code V15.82) or tobacco use disorder (ICD-9 305.1) before and after a cancer diagnosis. Use of smoking-cessation treatments before and after the cancer diagnosis was defined as documentation of smoking-cessation counseling (Healthcare Common Procedure Coding System codes 99406, 99407, G0375, and G0376) or pharmacotherapy (eg, nicotine replacement therapy, bupropion, varenicline).

RESULTS—From 2007 to 2011, among 23,703 patients in the KCR, we discerned a valid prediagnosis smoking status for 78%. KCR data only (72%), claims data only (6%), and a combination of both data sources (22%) were used to determine valid smoking status.

CORRESPONDING AUTHOR Michael Shayne Gallaway, PhD, MPH, Division of Cancer Prevention and Control; Centers for Disease Control and Prevention, 4770 Buford Hwy, MS F76, Atlanta, GA 30341; mgallaway@cdc.gov.

AUTHOR CONTRIBUTIONS

Conception and design: Michael Shayne Gallaway, Bin Huang, Tom Tucker, Eric Durbin, David Siegel
Collection and assembly of data: Michael Shayne Gallaway, Bin Huang, Jaclyn McDowell, Tom Tucker, Eric Durbin, David Siegel
Data analysis and interpretation: Michael Shayne Gallaway, Quan Chen, Tom Tucker, Eric Durbin, David Siegel, Eric Tai
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of work: All authors

PRIOR PRESENTATION

Presented at the North American Association of Central Cancer Registries meeting, Pittsburgh, PA, June 9–14, 2018.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. 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/rwcorascopubs.org/cci/author-center.

Approximately 4% of patients with cancer identified as smokers (n = 11,968) and were provided smoking-cessation counseling, and 3% were prescribed pharmacotherapy for smoking cessation.

CONCLUSION—Augmenting KCR data with medical claims data increased capture of smoking status and use of smoking-cessation modalities. Cancer registries interested in exploring smoking status to influence treatment and research activities could consider a similar approach, particularly if their registry does not capture smoking status for a majority of patients.

INTRODUCTION

Linkage of cancer registry data with complementary data sources can be an informative way to expand what is known about patients and their treatment, and aid improving delivery of care.[1–5] A report from the Institute of Medicine[6] recommended greater focus on linking data between registry and administrative data to broaden relevant measures for study. Cancer registries typically contain useful data on cancer severity, histology, and diagnosis, but often lack complete information on individual patient treatment, outcomes, and/or risk factors.[6] Medical claims–based data can add detailed information about the costs and use of medical services. Linkage between multiple complementary data sources can make it possible to explore new questions in large population-based patient samples, including relevant health care disparities. One example of a measure not typically captured by US cancer registries is smoking history. Some studies have explored the use of hospital records to identify smoking history of patients with cancer.[7,8]

Documenting smoking status at the time of a cancer diagnosis is important for helping patients with cancer who smoke quit, by assessing risks of multiple cancers and developing treatment and survivor care plans; and the availability of smoking status could further inform the investigation and/or evaluation of such, as has been demonstrated in a number of registries outside the United States.[8–11] Collection of smoking history for patients in cancer registries in the United States is nonstandardized, thus preventing large population-based studies from assessing the role of smoking on cancer treatment and outcomes. Cigarette smoking is harmful for people previously diagnosed with cancers,[12–15] and continued smoking negatively affects cancer treatment.[16–19] Tobacco cessation treatments can be effective among smokers with cancer.[20–22] Numerous methods are available to assist smokers with quitting, including counseling, pharmacotherapy, and nicotine replacement therapy (NRT). There is a lack of research on specific smoking-cessation modalities used among cancer survivors.[23]

The Kentucky Cancer Registry (KCR) conducted a data linkage to augment its cancer registry data with health claims data, such as that from Medicare, Medicaid, state employee insurance, and private insurance group claims. The intent was to improve the quality of the registry data and inform cancer care and outcomes research, while also providing an empirical basis to assess adherence to evidence-based quality-of-care measures and to compare patterns of care between Appalachian and non-Appalachian Kentucky populations. The purpose of this study was to explore whether patient smoking status and smoking-cessation modalities could be identified using the KCR augmented by linkage with health claims data.

METHODS

Data Sources

The KCR includes first, primary, invasive cancer diagnoses from 2007 to 2011. Cancers included lung and bronchus, colorectal, pancreatic, breast, ovary, and prostate, in accordance with SEER site recodes.[24] Patients in the registry captured through autopsy or death certificate only were excluded. Health claims data for patients with continuous enrollment coverage for 12 months before and 12 months after diagnosis included Medicare (SEER-Medicare, 2000 to 2012), Medicaid (Kentucky Family Health Service, 2000 to 2015), and private insurance groups (2006 to 2015).

Data Linkage

The linkage of SEER-Medicare data is a collaborative effort of the National Cancer Institute, the SEER registries, and the Centers for Medicare and Medicaid Services[25]; data for this study were obtained subsequent to linkage. Probabilistic data linkage was then performed using Link*Plus[26] to identify additional matches between the KCR and claims from Medicaid, state employee insurance, and private insurance groups. Linkage varied slightly depending on specific requirements and data use agreements. The direct method of probabilistic linkage was applied using blocking variables (social security number, birthdate, first name, middle name, last name) and matching variables (social security number, birthdate, first name, middle name, last name, sex) to calculate the m probability and u probability.[27] A minimum linkage value (ie, score indicating, for any pair of records, how likely it is that they both refer to the same person) of five or higher was considered a match. All potential matches were manually reviewed to verify true matches. Approximately 81.2% of cancer cases were linked with at least one claim.

Smoking Status

Smoking history was examined during three periods before and after a cancer diagnosis: 36 months before a cancer diagnosis, 12 months before a cancer diagnosis, and from 12 months before to 12 months after a cancer diagnosis. These periods were selected for comparison because of the likelihood of smoking behaviors before and after a cancer diagnosis,[15,19,22,23] and also as a means to verify the sensitivity to which differences may exist during these periods when using claims data to classify smoking and smoking-cessation history. Patients were classified with a history of smoking before the cancer diagnosis if documentation of personal history of tobacco use (International Classification of Diseases, Ninth Revision [ICD-9] code V15.82) or tobacco use disorder (ICD-9 305.1) was identified in linked records. Patients were classified with use of smoking-cessation treatments before and after the cancer diagnosis if documentation of smoking-cessation counseling (Healthcare Common Procedure Coding System codes 99406, 99407, G0375, and G0376) or pharmacotherapy (ie, NRT, bupropion, varenicline) was identified in linked records for the 12 months before or 12 months after diagnosis (or during the month of diagnosis) among patients classified as smokers.

Demographic Characteristics

We compared demographic characteristics among all patients by smoking status. Demographics assessed included sex, age (< 25, 25 to 44, 45 to 64, ≥ 65 years); marital status (married [married or living with partner], previously married [separated, divorced, widowed], never married, unknown); race/ethnicity (white non-Hispanic, black non-Hispanic, other [Asian, American Indian, Alaskan Native, multiple race, Hispanic {any race}]); education (percentage with high school education at county level, categorized by quartiles: very low, low, moderate, high); poverty (percentage below poverty line at county level, categorized by quartiles: low, moderate, high, very high); and health care insurance status (uninsured, private, Medicaid, Medicare, other public, unknown). In the United States, typically the elderly are covered by Medicare, low-income individuals are covered by Medicaid, most other people rely on their employer to provide health insurance, and a small number pay for their health care privately. Many others do not have any health care insurance. We also compared Appalachian status (ie, included or not included based on Appalachian Region Commission, https://www.arc.gov/appalachian_region/CountiesinAppalachia.asp). The Appalachian region of the eastern United States (including Kentucky) is home to more than 25 million people in mostly isolated mountainous areas, many of whom experience problems of rural poverty, and inadequate jobs, services, transportation, education, and infrastructure.

Statistical Analysis

The prevalence of smoking status (yes, no, unknown) was estimated for the KCR, medical claims, and combined linkage dataset; the prevalence of pre- and postdiagnosis smoking cessation was estimated for medical claims. We estimated the percent contribution of smoking status to the overall combined linkage for each data source (ie, KCR only, medical claims only, both KCR and medical claims, or neither). χ^2 test statistics were calculated with a threshold of $P < .05$ used as a measure of significant difference across relevant strata for patients with cancer with a valid smoking status. All analyses were completed using SAS Enterprise Guide, version 9.4 (SAS Institute, Cary, NC).

RESULTS

From 2007 to 2011, 23,703 patients were identified in the KCR along with the cancer sites. A history of smoking varied only slightly by period of assessment: 3 years before cancer diagnosis (53%), 1 year before cancer diagnosis (50%), or from 1 year before to 1 year after cancer diagnosis (55%) (Table 1). The contribution of claims data identifying smoking history was highest for 1 year before to 1 year after cancer diagnosis (40%). A history of smoking also varied by cancer type, with the highest prevalence observed among tobacco-associated cancers (ie, lung, pancreas, and colorectal).

Using the period 1 year before cancer diagnosis, we were able to discern a valid precancer diagnosis smoking status for 77% of patients (50% yes; 27% no). Smoking status was unknown for 22% of patients. Smoking status was identified for 18,388 patients in this population using KCR data only (72%), a combination of both data sources (22%), and claims data only (6%; Fig 1).

A history of smoking in the 12 months before a cancer diagnosis differed significantly by sex, age, marital status, race/ethnicity, education status, poverty level, Appalachian status, and insurance status. Smoking was more common among persons who were male (74%), aged 45 to 64 years (70%), previously married (65%), non-Hispanic black (69%), very low education (67%), of low (66%) or very high poverty (67%), of Appalachian status (66%), or who had Medicaid (79%) or unknown (83%) insurance (Table 2).

Smoking-cessation counseling and pharmacotherapy were noted in claims data before and after a primary cancer diagnosis among patients with cancer who were smokers (Table 3). Approximately 4% of patients with cancer identified as precancer diagnosis smokers (n = 11,968) were provided smoking-cessation counseling and 3% were prescribed cessation medications.

DISCUSSION

Augmenting the KCR with Medicare, Medicaid, and private insurance group claims data resulted in more fully capturing smoking status and use of smoking-cessation modalities within a population-based cancer registry for multiple types of cancers. Overall, smoking status was captured during the year before a cancer diagnosis for 77% of patients after linkage of the KCR and medical claims data; and each individual source of data contributed significantly. The linkage of the KCR with medical claims data yielded smoking status information for an additional 6% (n = 1,116 of 18,388) of the patient population. Although the linkage in this population ultimately may have resulted in the addition of smoking status data for a little more than 1,000 patients with cancer, claims data positively identified smoking status for a much higher percentage depending on the period of assessment. Thus, in a registry where smoking status is not collected or collected minimally, the potential linkage addition of valid patient smoking-status potential linkage addition of valid patient smoking-status information for the population could be significant. Therefore, linkage with claims data may be a useful resource for registries to obtain smoking status information.

This study demonstrated that a linkage between a population-based cancer registry and medical claims data successfully resulted in more fully capturing patient information related to prediagnosis smoking status and documentation of smoking-cessation counseling and medication. Precancer diagnosis smoking status in this patient population (50%) was higher than what has been observed in the general Kentucky population[28] but was consistent with estimates for prediagnosis smoking among patients with cancer.[7,29,30] We would also expect a higher percentage of the population diagnosed with cancers, specifically cancers causally linked with smoking (ie, lung, colorectal, pancreas), to be current or former smokers.[23,31] The percentage of patients (documented as smokers before cancer diagnosis) whose medical claims show provision of smoking-cessation counseling or pharmacotherapy may be an underrepresentation of patients who used cessation treatments (eg, not all providers will submit claims for provision of counseling services, NRT is also available over the counter). We were unable to determine the total percentage of patients who made an attempt to quit smoking, because the majority of cessation attempts are unassisted (ie, no counseling or pharmacotherapy) and thus not captured in medical claims.[32] However, assuming approximately one-half of documented smokers in this patient population (11,921 × 0.50 =

5,961) made a recent attempt to quit (consistent with recent research on quit attempts made[33]), the estimated use of smoking-cessation counseling (n = 427 of 5,961; 7.2%) is not significantly different from recent national self-reported estimates (7%).[33] However, the estimated use of smoking-cessation pharmacology in this population (n = 358 of 5,961; 6%) is considerably lower than reported for adult smokers during 2015, as reported by Babb et al (29%).[33]

Despite the stated limitations, the low rate of smoking-cessation counseling and NRT use among persons diagnosed with cancer in this population is noteworthy. A previous study determined that even after a cancer diagnosis, approximately one in eight cancer survivors continued to smoke.[34] Continued smoking after a cancer diagnosis is associated with adverse health outcomes, including a higher risk of death, negative effects on cancer treatment, and increased risk of another cancer. However, all cancer survivors may not be aware of how continued smoking after a cancer diagnosis negatively affects their health. All smokers should receive advice to quit smoking by a health professional, but approximately one-third do not.[34] Health professionals could consistently advise cancer survivors about the increased risks associated with continued smoking, provide them with cessation counseling and medications, refer them to other free cessation resources, and inform them of cessation treatments covered by their health insurance. To increase cessation pharmacology use in the population of patients with cancer, state tobacco-control programs as well as hospitals, clinics, and quitlines could partner with cancer treatment centers to build and implement sustainable tobacco-cessation treatment programs to routinely address tobacco cessation with patients with cancer.

Linking relevant complementary data sources can expand availability and utility of data variables that can add additional context to investigations and evaluations of populations of patients with cancer. A few population-based registries in the United States[7,35,36] and outside the United States[8,9,11] have used patient medical records to more fully capture patients' smoking status. Using this approach, information about smoking was captured for between 43% and 92% of patients.[7–9,11,36] Routinely abstracting information on smoking status from electronic medical records (EMRs) can be useful, even if the data are not complete.[8] However, depending on the EMR, this can require extensive time and effort. Our goal was to develop a sustainable Kentucky Cancer Quality Outcome Research Data System that would contain a range of contextual data variables while minimizing the amount of labor needed to capture additional patient information.[37] Direct linkage of the KCR to Medicare, Medicaid, and private insurance group claims data resulted in augmentation of smoking-status information. However, the usefulness of this procedure to a particular registry or population will likely vary depending on existing efforts to capture this information, such as use of EMRs.

Linking to additional patient risk-factor data existing in medical claims is in line with priorities established by the Institute of Medicine for comparative effectiveness research[38] aimed at assisting consumers, clinicians, purchasers, and policy makers to make informed decisions to improve individual and population health. The Centers for Disease Control and Prevention established 10 specialized registries within the National Program of Cancer Registries to expand data collection to include additional data variables, such as smoking

status, for all cancers.[39] Linkages were performed to enhance registry data with census data, National Death Index files, hospital discharge data, and each state's breast and cervical early-detection programs, resulting in the capture of cigarette-smoking information for approximately 66% of patients during 2011 and 51% between 2011 and 2013.[35,39] Continued expansion of state and national registry data beyond typical measures (eg, stage, histology, diagnosis) to include more patient-centered information has the potential to influence a variety of treatment and research activities.

The use of medical claims records to identify smoking history and receipt of smoking-cessation treatment has limitations. Although the use of ICD-9 tobacco-use codes is a reliable method for identifying smoking status,[40] the absence of an ICD-9 code does not indicate a lack of smoking. It is also possible that although a patient is listed as a smoker, they may have quit since their smoking status was last documented. In this study, we were unable to differentiate the type of tobacco used, although a previous study examining population-based cancer registries, using enhanced data collection to ascertain tobacco use, [34] showed that only a small percentage of people (2%) used nonsmoked tobacco products, and it has been estimated nationally that approximately 3.5% of adults use these products regularly.[41] The existence of relevant codes for smoking-cessation counseling and pharmacology may not accurately reflect whether a registry patient actually adhered to such treatment or filled the prescription.[42] Furthermore, it is possible that not all prescription medication use is accurately reflected in pharmacy claims data,[43,44] as a result of filling prescriptions through various benefit programs (eg, dual eligibility, spousal benefits).[42,44] Finally, medical claims used to approximate receipt of smoking-cessation treatment are limited to physician-provided or -prescribed counseling and/or pharmacology, and does not include free NRT received from state quitlines, prescription samples obtained from physicians' offices,[42] or some of the most common methods used to stop smoking (eg, unassisted, nonprescribed, or nonreimbursed over-the-counter smoking-cessation aids).[32]

Documentation of smoking status history at the time of a cancer diagnosis is important for helping patients with cancer who smoke quit, by assessing future risks of multiple cancers and for developing treatment and survivor care plans. Augmenting the KCR through linkage with medical claims data resulted in more fully capturing smoking status and use of smoking-cessation medications and counseling within a population-based cancer population. Tobacco-control programs may be able to leverage these data to implement tobacco-cessation programs to prevent tobacco use, promote cessation,[45] identify and address disparities related to tobacco use and cancer outcomes among different population groups. Cancer registries considering expanding available measures beyond cancer stage, histology, and diagnosis to inform treatment and research activities could consider a similar approach. In addition to being used to track smoking status and receipt of smoking-cessation treatments, medical claims data may be a critical data source for identifying other modifiable risk factors within cancer populations.

Acknowledgments

SUPPORT

Supported by the Centers for Disease Control and Prevention (Grant No. U48DP0085014–01 SIP =14–017; B.H.) and the National Cancer Institute (Grant No. P30 CA177558; T.T.).

The findings and conclusions in this report are those of the authors and do not necessarily reflect the official position of the Centers for Disease Control and Prevention or The University of Kentucky.

REFERENCES

1. Bradley CJ, Given CW, Roberts C: Late stage cancers in a Medicaid-insured population. Med Care 41:722–728, 2003 [PubMed: 12773838]
2. Ganz PA, Desmond KA, Leedham B, et al.: Quality of life in long-term, disease-free survivors of breast cancer: A follow-up study. J Natl Cancer Inst 94:39–49, 2002 [PubMed: 11773281]
3. Ganz PA, Greendale GA, Petersen L, et al.: Breast cancer in younger women: Reproductive and late health effects of treatment. J Clin Oncol 21:4184–4193, 2003 [PubMed: 14615446]
4. Koroukian SM: Linking the Ohio Cancer Incidence Surveillance System with Medicare, Medicaid, and clinical data from home health care and long term care assessment instruments: Paving the way for new research endeavors in geriatric oncology. J Registry Manag 35:156–165, 2008 [PubMed: 25132914]
5. Hewitt M, Simon JV (eds): Enhancing Data Systems to Improve the Quality of Cancer Care Washington, DC, National Academies Press, 2000
6. Hewitt M, Greenfield S, Stovall E (eds): From Cancer Patient to Cancer Survivor: Lost in Translation Washington, DC, National Academies Press, 2006
7. Polednak AP: Obtaining smoking histories for population-based studies on multiple primary cancers: Connecticut, 2002. Int J Cancer 119:233–235, 2006 [PubMed: 16432840]
8. Prochazka M, Hall P, Gagliardi G, et al.: Ionizing radiation and tobacco use increases the risk of a subsequent lung carcinoma in women with breast cancer: Case-only design. J Clin Oncol 23:7467–7474, 2005 [PubMed: 16234513]
9. Sharp L, McDevitt J, Brown C, et al.: Association between smoking at diagnosis and cause-specific survival in patients with rectal cancer: Results from a population-based analysis of 10,794 cases. Cancer 123:2543–2550, 2017 [PubMed: 28297071]
10. Radzikowska E, Głaz P, Roszkowski K: Lung cancer in women: Age, smoking, histology, performance status, stage, initial treatment and survival. Population-based study of 20 561 cases. Ann Oncol 13:1087–1093, 2002 [PubMed: 12176788]
11. Sifaki-Pistolla D, Lionis C, Georgoulias V, et al.: Lung cancer and tobacco smoking in Crete, Greece: Reflections from a population-based cancer registry from 1992 to 2013. Tob Induc Dis 15:6, 2017 [PubMed: 28123354]
12. Klosky JL, Tyc VL, Garces-Webb DM, et al.: Emerging issues in smoking among adolescent and adult cancer survivors: A comprehensive review. Cancer 110:2408–2419, 2007 [PubMed: 17932906]
13. Mackenbach JP, Borsboom GJ, Nusselder WJ, et al.: Determinants of levels and changes of physical functioning in chronically ill persons: Results from the GLOBE Study. J Epidemiol Community Health 55:631–638, 2001 [PubMed: 11511641]
14. Mariotto AB, Rowland JH, Ries LA, et al.: Multiple cancer prevalence: A growing challenge in long-term survivorship. Cancer Epidemiol Biomarkers Prev 16:566–571, 2007 [PubMed: 17372253]
15. Richardson GE, Tucker MA, Venzon DJ, et al.: Smoking cessation after successful treatment of small-cell lung cancer is associated with fewer smoking-related second primary cancers. Ann Intern Med 119:383–390, 1993 [PubMed: 8393311]
16. Browman GP, Wong G, Hodson I, et al.: Influence of cigarette smoking on the efficacy of radiation therapy in head and neck cancer. N Engl J Med 328:159–163, 1993 [PubMed: 8417381]
17. Des Rochers C, Dische S, Saunders MI: The problem of cigarette smoking in radiotherapy for cancer in the head and neck. Clin Oncol (R Coll Radiol) 4:214–216, 1992 [PubMed: 1622882]
18. Gritz ER, Dresler C, Sarna L: Smoking, the missing drug interaction in clinical trials: Ignoring the obvious. Cancer Epidemiol Biomarkers Prev 14:2287–2293, 2005 [PubMed: 16214906]

19. Mason DP, Subramanian S, Nowicki ER, et al.: Impact of smoking cessation before resection of lung cancer: A Society of Thoracic Surgeons General Thoracic Surgery Database study. Ann Thorac Surg 88:362–370, 2009 [PubMed: 19632374]
20. Kawahara M, Ushijima S, Kamimori T, et al.: Second primary tumours in more than 2-year disease-free survivors of small-cell lung cancer in Japan: The role of smoking cessation. Br J Cancer 78:409–412, 1998 [PubMed: 9703291]
21. Tucker MA, Murray N, Shaw EG, et al.: Second primary cancers related to smoking and treatment of small-cell lung cancer. Lung Cancer Working Cadre. J Natl Cancer Inst 89:1782–1788, 1997 [PubMed: 9392619]
22. Tobacco Use and Dependence Guideline Panel: Treating tobacco use and dependence: 2008 Update Rockville, MD, US Department of Health and Human Services, 2008
23. Underwood JM, Townsend JS, Tai E, et al.: Persistent cigarette smoking and other tobacco use after a tobacco-related cancer diagnosis. J Cancer Surviv 6:333–344, 2012 [PubMed: 22706885]
24. National Cancer Institute: Surveillance, Epidemiology, and End Results Program, Site Recode ICD-O-3/WHO 2008 Definition, 2008 https://seer.cancer.gov/siterecode/icdo3_dwhoheme/
25. 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 40:IV-3–IV-18, 2002 (8, Suppl)
26. Centers for Disease Control and Prevention: Registry Plus Link Plus, 2015 https://www.cdc.gov/cancer/npcr/tools/registryplus/lp.htm
27. Dusetzina SB, Tyree S, Meyer AM, et al.: Linking Data for Health Services Research: A Framework and Instructional Guide Rockville, MD, AHRQ Methods for Effective Health Care,2014
28. Nguyen KH, Marshall L, Brown S, et al.: State-specific prevalence of current cigarette smoking and smokeless tobacco use among adults - United States, 2014. MMWR Morb Mortal Wkly Rep 65:1045–1051, 2016 [PubMed: 27711031]
29. Shiels MS, Gibson T, Sampson J, et al.: Cigarette smoking prior to first cancer and risk of second smoking-associated cancers among survivors of bladder, kidney, head and neck, and stage I lung cancers. J Clin Oncol 32:3989–3995, 2014 [PubMed: 25385740]
30. Passarelli MN, Newcomb PA, Hampton JM, et al.: Cigarette smoking before and after breast cancer diagnosis: Mortality from breast cancer and smoking-related diseases. J Clin Oncol 34:1315–1322, 2016 [PubMed: 26811527]
31. US Department of Health and Human Services: The Health Consequences of Smoking: 50 Years of Progress: A Report of the Surgeon General Atlanta, GA, US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014
32. Caraballo RS, Shafer PR, Patel D, et al.: Quit methods used by US adult cigarette smokers, 2014–2016. Prev Chronic Dis 14:E32, 2017 [PubMed: 28409740]
33. Babb S, Malarcher A, Schauer G, et al.: Quitting smoking among adults — United States, 2000–2015. MMWR Morb Mortal Wkly Rep 65:1457–1464, 2017 [PubMed: 28056007]
34. Gallaway MS, Glover-Kudon R, Momin B, et al.: Smoking cessation attitudes and practices among cancer survivors - United States, 2015. J Cancer Surv 13:66–74, 2019
35. Knowlton R, Gershman S, Solis A, et al.: An assessment of the reliability of race, Hispanic ethnicity, birthplace, and tobacco history data in the Massachusetts cancer registry, 2005–2009. J Registry Manag 41:146–150, 2014 [PubMed: 25419609]
36. Siegel DA, Henley SJ, Wike JM, et al.: Capture of tobacco use among population-based registries: Findings from 10 National Program of Cancer Registries states. Cancer 124:2381–2389, 2018 [PubMed: 29579317]
37. University of Kentucky: Kentucky Cancer Registry. The population-based central registry for the Commonwealth of Kentucky http://www.kcr.uky.edu/research/types.php
38. Institute of Medicine Committee on Comparative Effectiveness Research Prioritization Board of Health Care Services: Initial National Priorities for Comparative Effectiveness Research Washington, DC, National Academies Press, 1999

39. Chen VW, Eheman CR, Johnson CJ, et al.: Enhancing cancer registry data for comparative effectiveness research (CER) project: Overview and methodology. J Registry Manag 41:103–112, 2014 [PubMed: 25419602]

40. Wiley LK, Shah A, Xu H, et al.: ICD-9 tobacco use codes are effective identifiers of smoking status. J Am Med Inform Assoc 20:652–658, 2013 [PubMed: 23396545]

41. Hu SS, Neff L, Agaku IT, et al.: Tobacco product use among adults - United States, 2013–2014. MMWR Morb Mortal Wkly Rep 65:685–691, 2016 [PubMed: 27416365]

42. Crystal S, Akincigil A, Bilder S, et al.: Studying prescription drug use and outcomes with medicaid claims data: strengths, limitations, and strategies. Med Care 45:S58–S65, 2007 (10, Suppl 2) [PubMed: 17909385]

43. Lauffenburger JC, Balasubramanian A, Farley JF, et al.: Completeness of prescription information in US commercial claims databases. Pharmacoepidemiol Drug Saf 22:899–906, 2013 [PubMed: 23696101]

44. Cepeda MS, Fife D, Denarié M, et al.: Quantification of missing prescriptions in commercial claims databases: results of a cohort study. Pharmacoepidemiol Drug Saf 26:386–392, 2017 [PubMed: 28120552]

45. Krebs P, Rogers E, Wong H, et al.: Assessing the Accuracy of Registry Based Tobacco Use Status and Utility for Patient Recruitment Into Tobacco Trials Albuquerque, NM, NAACR, 2017

CONTEXT

Key Objective

How can cancer registry data be linked with health claims data to increase what is known about patient smoking status and smoking cessation?

Knowledge Generated

Smoking status and smoking cessation were captured in 22% of health claims data. Smoking cessation was present for less than 5% of patients after cancer diagnosis.

Relevance

Cancer registries that do not capture smoking status could use claims data to identify what is known about these behaviors. Health professionals should consistently advise patients with cancer about the increased risks associated with continued smoking, provide them with cessation counseling and medications, refer them to other free cessation resources, and inform them of cessation treatments covered by their health insurance.

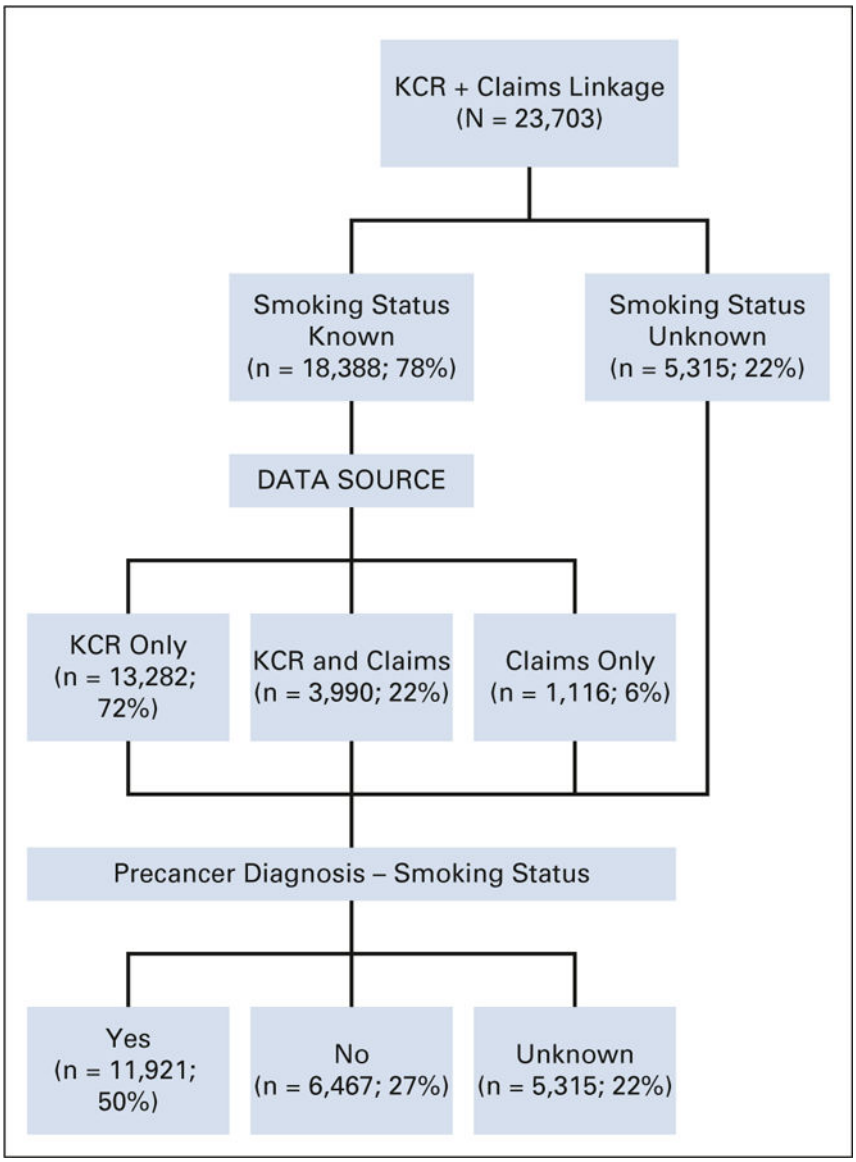

FIG 1.

Kentucky Cancer Registry (KCR) and medical claims contribution to smoking status capture for the 12 months before cancer diagnosis.

TABLE 1.

Comparison of Smoking History During Varying Periods Before and After Cancer Diagnosis Overall and by Cancer Type (N = 23,703)

		Smoking History					
		3 Years Before Cancer Diagnosis*		1 Year Before Cancer Diagnosis†		From 1 Year Before to 1 Year After Cancer Diagnosis‡	
Cancer Type	Total No. of Patients	No.	%	No.	%	No.	%
Total	23,703						
Augmented		12,457	52.6	11,921	50.3	13,092	55.2
KCR		10,746	45.3	10,746	45.3	10,728	45.3
Claims		6,936	29.3	5,229	22.1	9,456	39.9
Lung and bronchus	5,096						
Augmented		4,379	85.9	4,287	84.1	4,505	88.4
KCR		3,924	77.0	3,924	77.0	3,924	77.0
Claims		3,053	59.9	2,555	50.1	3,824	75.0
Colorectal	4,563						
Augmented		2,085	45.7	1,986	43.5	2,228	48.8
KCR		1,762	38.6	1,762	38.6	1,762	38.6
Claims		1,108	24.3	843	18.5	1,554	34.1
Pancreatic	374						
Augmented		191	51.1	180	48.1	207	55.3
KCR		143	38.2	143	38.2	143	38.2
Claims		130	34.8	107	28.6	170	45.5
Breast	13,670						
Augmented		2,241	37.9	2,148	36.3	2,413	40.8
KCR		1,986	33.6	1,986	33.6	1,986	33.6
Claims		1,047	17.7	690	11.7	1,694	28.6
Ovarian	412						
Augmented		139	33.7	132	32.0	148	35.9
KCR		112	27.2	112	27.2	112	27.2
Claims		81	19.7	69	16.7	111	26.9

JCO Clin Cancer Inform. Author manuscript; available in PMC 2020 May 01.

		Smoking History					
		3 Years Before Cancer Diagnosis*		1 Year Before Cancer Diagnosis†		From 1 Year Before to 1 Year After Cancer Diagnosis‡	
Cancer Type	Total No. of Patients	No.	%	No.	%	No.	%
Prostate	7,342						
Augmented		3,419	46.6	3,188	43.4	3,632	49.5
KCR		2,819	38.4	2,819	38.4	2,819	38.4
Claims		1,496	20.4	965	13.1	2,141	29.2

Abbreviation: KCR, Kentucky Cancer Registry.

*Documentation of personal history of tobacco use (International Classification of Diseases, Ninth Revision [ICD-9] V15.82) or tobacco use disorder (ICD-9 305.1) was identified in KCR or claims records for the 36 months before the cancer diagnosis.

†Documentation of personal history of tobacco use (ICD-9 V15.82) or tobacco use disorder (ICD-9 305.1) was identified in KCR or claims records for the 12 months before the cancer diagnosis.

‡Documentation of personal history of tobacco use (ICD-9 V15.82) or tobacco use disorder (ICD-9 305.1) was identified in KCR or claims records for the 12 months before the cancer diagnosis or for the 12 months after the cancer diagnosis.

TABLE 2.

Smoking History and Demographic Characteristics for Patient Population (n = 18,388)

	Smoking History*				
	No		Yes		
Characteristic	**No.**	**%**	**No.**	**%**	***p*†**
Overall	6,467	35.2	11,921	64.8	
Sex					< .001
Male	2,347	26.0	6,695	74.0	
Female	4,120	44.1	5,226	55.9	
Age, years					< .001
< 25	15	75.0	5	25.0	
25–44	183	46.4	211	53.6	
45–64	1,336	29.8	3,148	70.2	
≥ 65	4,933	36.6	8,557	63.4	
Marital status					< .01
Married	3,785	35.8	6,790	64.2	
Previously married	2,051	34.9	3,828	65.1	
Never married	500	38.1	814	61.9	
Unknown	131	21.1	489	78.9	
Race/ethnicity					0.02
White, non-Hispanic	5,952	35.4	10,859	64.6	
Black, non-Hispanic	352	31.5	764	68.5	
Hispanic	27	45.8	32	54.2	
Other	136	33.8	266	66.2	
Education‡					< .001§
Very low	1,485	32.7	3,060	67.3	
Low	1,616	36.0	2,876	64.0	
Moderate	2,520	34.9	4,700	65.1	
High	846	39.7	1,285	60.3	
Poverty¶					< .001§
Low	1,538	33.6	3,040	66.4	
Moderate	1764	34.8	3,305	65.2	
High	1,701	39.1	2,650	60.9	
Very high	1,464	33.3	2,926	66.7	
Appalachian status					< .01
Appalachian	1,813	33.7	3,561	66.3	
Non-Appalachian	4,654	35.8	8,360	64.2	
Insurance					< .001
Not insured	15	27.8	39	72.2	

	Smoking History*				
	No		Yes		
Characteristic	**No.**	**%**	**No.**	**%**	***P*†**
Private	1,345	42.8	1,799	57.2	
Medicaid	212	21.1	793	78.9	
Medicare	4,823	34.9	8,981	65.1	
Other public	34	22.2	119	77.8	
Unknown	38	16.9	187	83.1	

*Personal history of tobacco use (International Classification of Diseases, Ninth Revision [ICD-9] code V15.82) or tobacco use disorder (ICD-9 305.1) identified in linked records for the 12 months before the cancer diagnosis.

†χ^2 test.

‡Percentage with high school education at county level, categorized by listed quartiles.

§Percentage below poverty line at county level, categorized by listed quartiles.

¶*P* value for linear trend was not statistically significant.

TABLE 3.

Pre- and Postcancer Diagnosis Smoking Cessation Counseling or Pharmacotherapy Status Identified for Kentucky Cancer Registry Patient Population Smokers (n = 11,968)

	Smoking Cessation Treatment							
	None		Any		Prediagnosis*		Postdiagnosis†	
Type of Cessation Treatment	No.	%	No.	%	No.	%	No.	%
Counseling (tobacco use)‡	11,494	96.0	427	3.6	195	1.6	278	2.3
Pharmacotherapy§	11,563	96.6	358	3.0	207	1.7	259	2.2

*Twelve months before cancer diagnosis.

†Twelve months after diagnosis (or during the month of diagnosis).

‡Documentation of smoking cessation counseling (Healthcare Common Procedure Coding System codes: 99406, 99407, G0375, G0376).

§Documentation of pharmacotherapy (eg, nicotine replacement therapy, bupropion, varenicline).

JCO Clin Cancer Inform. Author manuscript; available in PMC 2020 May 01.