



Published in final edited form as:

*Cancer*. 2018 June 01; 124(11): 2381–2389. doi:10.1002/cncr.31326.

## Capture of Tobacco Use Among Population-Based Registries: Findings From 10 National Program of Cancer Registries States

David A. Siegel, MD, MPH<sup>1,2</sup>, S. Jane Henley, MSPH<sup>2</sup>, Jennifer M. Wike, MBA, MPH<sup>3</sup>, A. Blythe Ryerson, PhD, MPH<sup>2</sup>, Christopher J. Johnson, MPH<sup>4</sup>, Judy R. Rees, BM, BCh, MPH, PhD<sup>5,6</sup>, Lori A. Pollack, MD, MPH<sup>2</sup>, and Enhancement of NPCR for Comparative Effectiveness Research Team

<sup>1</sup>Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, Georgia

<sup>2</sup>Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia

<sup>3</sup>DB Consulting Group, Inc, CDC-NPCR Contractor, Atlanta, Georgia

<sup>4</sup>Cancer Data Registry of Idaho, Boise, Idaho

<sup>5</sup>New Hampshire State Cancer Registry, Lebanon, New Hampshire

<sup>6</sup>Department of Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire

### Abstract

**BACKGROUND:** Tobacco use data are important when the epidemiology and prognosis of tobacco-associated cancers are being defined. Central cancer registries in 10 National Program of Cancer Registries states pilot-tested the collection of standardized tobacco use variables. This study evaluated the capture of tobacco use data and examined smoking prevalence among cancer patients.

**METHODS:** Participating registries collected data about the use of tobacco—cigarettes, other smoked tobacco, and smokeless tobacco—for cases diagnosed during 2011-2013. The percentage

---

**Corresponding author:** David A. Siegel, MD, MPH, Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Highway Northeast, Mailstop F-76, Atlanta, GA 30341; dsiegel@cdc.gov.

#### AUTHOR CONTRIBUTIONS

**David A. Siegel:** Conceptualization, formal analysis, investigation, methodology, visualization, writing—original draft, and writing—review and editing. **S. Jane Henley:** Conceptualization, investigation, methodology, and writing—review and editing. **Jennifer M. Wike:** Conceptualization, formal analysis, investigation, methodology, project administration, and writing—review and editing. **A. Blythe Ryerson:** Conceptualization, investigation, methodology, project administration, and writing—review and editing. **Christopher J. Johnson:** Conceptualization, data curation, investigation, project administration, validation, and writing—review and editing. **Judy R. Rees:** Conceptualization, data curation, investigation, project administration, validation, and writing—review and editing. **Lori A. Pollack:** Conceptualization, formal analysis, investigation, methodology, project administration, supervision, visualization, writing—original draft, and writing—review and editing.

#### CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

This article has been contributed to by US Government employees and their work is in the public domain in the USA.

The findings and conclusions are those of the authors and do not necessarily represent the official position of their affiliations or the Centers for Disease Control and Prevention.

of cases with known tobacco variable values was calculated, and the prevalence of tobacco use was analyzed by the primary cancer site and state.

**RESULTS:** Among 1,646,505 incident cancer cases, 51% had known cigarette use data: 18% were current users, 31% were former users, and 51% reported never using. The percentage of cases with a known status for both other smoked tobacco and smokeless tobacco was 43%, with 97% and 98% coded as never users, respectively. The percent known for cigarette use ranged from 27% to 81% by state and improved from 47% in 2011 to 59% in 2013 for all 10 states combined. The percent known for cigarette use and the prevalence of ever smoking cigarettes were highest for laryngeal cancer and tracheal, lung, and bronchus cancer.

**CONCLUSIONS:** Cancer registrars ascertained cigarette use for slightly more than half of all new cancer cases, but other tobacco-related fields were less complete. Studies to evaluate the validity of specific tobacco-related variables and the ability of cancer registries to capture this information from the medical record are needed to gauge the usefulness of collecting these variables through cancer surveillance systems.

### Keywords

cancer; epidemiology; registries; smoking; surveillance; tobacco

---

## INTRODUCTION

Tobacco use in the United States accounts for an estimated 167,805 cancer deaths each year.<sup>1</sup> Tobacco use at cancer diagnosis and during cancer care is associated with a worse disease prognosis.<sup>2-6</sup> US statistics on tobacco use are generated through surveys of the general population, such as the National Health Interview Survey (NHIS) and the Behavioral Risk Factor Surveillance System (BRFSS), and US cancer statistics are obtained from cancer patient medical records through cancer registries funded by the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC) or the Surveillance, Epidemiology, and End Results program of the National Cancer Institute.<sup>7</sup> Surveys that ask about tobacco use (eg, BRFSS) assess responses from patients with and without self-reported cancer diagnoses but do not link tobacco use with confirmed cancer diagnosis information from the medical record.

Using cancer registries to collect tobacco use data has the potential to generate more reliable data regarding tobacco use patterns for patients diagnosed with cancer. Tobacco use data at diagnosis may help health professionals to better understand how tobacco use affects a cancer prognosis,<sup>8</sup> and this in turn may inform patient counseling regarding risks associated with tobacco use as well as the management of survivorship care. In addition, tobacco use data collected in cancer registries could help health professionals to identify cancer patients for tobacco cessation interventions.

Many statewide individual central cancer registries collect tobacco use data, but these variables are not standardized among cancer registries. To improve the data-capture capabilities of cancer registries, in 2011, the CDC started an enhanced data collection project in comparative effectiveness research (CER), which was funded through the

American Recovery and Reinvestment Act.<sup>9</sup> The CER project secured funding for 10 statewide central cancer registries to collect and report new, standardized data variables, including tobacco use, height, and weight at diagnosis.<sup>9</sup> The quality and content of the tobacco use data from the CER project have not previously been reported. The purpose of this study was to evaluate the data quality of tobacco use data collected during 2011-2013 by the 10 CER states to inform recommendations on how best to collect these data in NPCR registries.

## MATERIALS AND METHODS

Data were analyzed from 10 state-based central cancer registries funded by CDC's NPCR program for the CER project, as detailed by Chen et al.<sup>9</sup> State registries included the registries for Alaska, California, Colorado, Florida, Idaho, Louisiana, North Carolina, New Hampshire, Rhode Island, and Texas.<sup>9</sup> The collection of tobacco use variables was required statewide for all patients diagnosed with cancers of any anatomic site during 2011-2013. NPCR data from the 10 CER states during 2011-2013 met US Cancer Statistics publication criteria and covered 27.3% of the US population.<sup>9</sup> The CER project was approved by the CDC institutional review board. Patient consent was not needed for this study because the submitted data were de-identified before they were received by the CDC.

Cases included all new diagnoses of cancer, including all behavior codes (benign, borderline, in situ, and malignant), and first and subsequent malignant neoplasms. The primary site was classified with site codes based on the *International Classification of Diseases for Oncology, Third Edition (ICD-O-3)*.<sup>10</sup> Further analysis was performed for tobacco-associated cancers, which are those cancers considered by the US surgeon general to be causally associated with tobacco use.<sup>11</sup> Tobacco-associated cancer cases<sup>11,12</sup> included *ICD-O-3* histology codes 8000 to 9049, 9056 to 9139, and 9141 to 9589. Sites included the lip, oral cavity, and pharynx (C000-C148); esophagus (C150-C159); stomach (C160-C169); colon and rectum (C180-C209 and C260); liver (C220); pancreas (C250-C259); larynx (C320-C329); trachea, lung, and bronchus (C339-C349); cervix (C530-C539); kidney and renal pelvis (C649-C659); urinary bladder (C670-C679); and acute myeloid leukemia (9840, 9861, 9865-9867, 9869, 9871-9874, 9895-9898, 9910-9911, and 9920).<sup>10,13</sup>

There were 4 tobacco use variables included in the CER data set and this analysis: 1) cigarette smoking, 2) other smoked tobacco (eg, pipes, cigars, and kreteks), 3) smokeless tobacco, and 4) tobacco use not otherwise specified (NOS). The use status for each tobacco question was coded as 1) never user, 2) current user, 3) former user who had quit within 1 year from the diagnosis, 4) former user who had quit more than 1 year before the diagnosis, 5) former user with an unknown time since quitting, or 6) unknown. These 4 tobacco use variables were developed specifically for the CER project by a multidisciplinary study team that included expertise in oncology, registry operations, cancer surveillance, and epidemiology and included physicians, researchers, and other public health professionals at the CDC and at participating states. The tobacco use variables were designed to capture data likely to be available for abstraction from medical records. The study team aimed to specify the type of tobacco used, to achieve mutually exclusive coding options for the type of tobacco use and status, and to create variables similar to historical tobacco use variables

already being collected by some states. The study team completed a literature review of previous data collection and epidemiological studies involving tobacco use variables to help inform the design of the data collection instrument. Once draft data collection fields were developed, the study team shared the design with the registry teams collecting the data to ensure that content was consistent and understood. Each state used the same single-field and interfiled edits and data submission record layout before data submission to ensure data quality and consistency.<sup>9</sup>

Cancer registry data were abstracted from information documented in the medical records. The sources of these data included nursing assessment notes, flow charts, vital stats, and other available sources from the patient's hospital or physician office medical record.<sup>14</sup> Data from the CER project did not include where tobacco use information was found within the medical chart. The tobacco use status was defined by the use status at the date of diagnosis, as abstracted from medical records. The percent known was calculated by the division of the number of cases coded with known values (with those coded as unknown being excluded) by the total number of cases, as previously done.<sup>15</sup> All analyses were conducted with SEER\*Stat version 8.3.2 (<https://seer.cancer.gov/seerstat/>).

## RESULTS

From 2011 to 2013, 1,646,505 new diagnoses of cancer were reported to the 10 CER registries. The majority of cases (50.5%) represented individuals aged 60 to 79 years, and the median age was 66 years (Table 1). The most common cancer diagnoses were digestive system cancer (16.8%), breast cancer (16.3%), and respiratory system cancer (12.5%). The study included 8846 cases (0.5%) from Alaska, 540,947 (32.9%) from California, 74,100 (4.5%) from Colorado, 369,638 (22.4%) from Florida, 24,994 (1.5%) from Idaho, 79,478 (4.8%) from Louisiana, 167,566 (10.2%) from North Carolina, 26,403 (1.6%) from New Hampshire, 20,240 (1.2%) from Rhode Island, and 334,293 (20.3%) from Texas.

The average percent known for tobacco use over the study interval was 51.4% (846,212 of 1,646,502) for cigarette use, 42.9% (705,639 of 1,646,502) for other smoked tobacco, 43.1% (709,635 of 1,646,502) for smokeless tobacco, and 43.4% (714,844 of 1,646,501) for tobacco use NOS. Among all tobacco use variables, the percent known was higher in females than males, was higher in blacks than all other races, and was higher in non-Hispanics than Hispanics. The percent known for cigarette use was highest for cases diagnosed at the ages of 40 to 79 years and was lowest for those aged 80 years. For all tobacco use variables, the percent known was highest for cases reported from a hospital or from a radiation or oncology center, which represented 88.8% and 1.9% of all cases, respectively.

For all 4 tobacco use variables, the percent known improved from 2011 to 2013 during the study interval (Fig. 1). The percent known varied by state and ranged from 26.7% to 81.3% for the cigarette use variable (Fig. 2).

Cigarette smoking was the most common form of tobacco use reported: 48.9% of cases ever smoked cigarettes, 2.8% ever smoked other tobacco products, and 1.8% ever used smokeless

tobacco products (Table 2). Current cigarette use was reported by 17.6% of cases, and former use was reported by 31.3%, with most of those users quitting more than 1 year before their diagnosis (Table 2).

Cigarette use had a higher percent known among cases diagnosed with a tobacco-associated cancer (55.3%) in comparison with cases diagnosed with cancers not associated with tobacco use (49.0%; Table 3). The prevalence of ever smoking cigarettes was also higher among cases diagnosed with a tobacco-associated cancer (63.2%) than other cases (39.0%; Table 3). The percent known for cigarette use and the prevalence of ever smoking cigarettes were highest among cases diagnosed with laryngeal cancer or tracheal, lung, or bronchus cancer (Table 3).

## DISCUSSION

This is the largest study to assess the capture of tobacco use by cancer registry record abstraction to date. Information about cigarette use was ascertained for slightly more than half of all new cancer cases in our study, but other tobacco use variables had lower percentages known. The percent known improved steadily between 2011 and 2013 for each tobacco use variable, but it showed wide variation by state. The percentage of current and former use of tobacco products other than cigarettes was <3% in each case. Tobacco-associated cancers had both a higher cigarette use percent known and higher recorded percentages for current and former cigarette users.

The percent known for tobacco use variables collected by cancer registries is not well described in the literature. A recent study by Sharp et al<sup>6</sup> using the National Cancer Registry Ireland reported a known smoking status for 77% of 10,794 rectal cancer cases. Using a variable vocabulary different from that of our study, the Massachusetts Cancer Registry, as part of a CDC-funded project from 2005 to 2009, reported 89.6% as the percent known for tobacco use variables among 5348 breast and colorectal cancer cases.<sup>16</sup> That study was limited to colorectal and breast cancers because they are more likely to be diagnosed and treated in a hospital setting. For the cigarette use variable, our study found 56.5% known for colorectal cancer and 56.0% known for breast cancer; these figures increased to 58.1% and 56.4%, respectively, when they were limited to cases found in hospital settings only. The percent known for cigarette use in our study may be higher than that for other tobacco use variables because cigarette use may be more often recorded in the medical record or may be easier for registrars to locate in the record. It is possible that the increase in the percent known from 2011 to 2013 reflects increased tobacco use reporting in the record or increased registrar familiarity with these registry variables.

The percent known for tobacco use variables in our study varied with several factors, including sex, age, race, ethnicity, state, and reporting source. These factors may have influenced a health care provider's likelihood to ask and record tobacco use, a patient's likelihood to report it, and the registrars' awareness of tracking and recording these variables. The wide variation in the percent known between states (26.7%-81.3% for the cigarette use variable) may reflect whether tobacco variables were already being collected before the CER study; differences in the availability, use, and comprehensiveness of

electronic health record (EHR) systems; or differences in variable acceptability. In addition, variable acceptability may be dependent on the emphasis placed by local and state registry leadership on recording tobacco use. Differences in coding practices by state, such as coding “never used” instead of “unknown,” or differences in the use of the tobacco use NOS code may explain further variation.

The percent known for cigarette use was higher for tobacco-associated cancers than other cancers and was highest for laryngeal cancer and tracheal, lung, and bronchus cancer. The higher percent known among tobacco-associated cancers may reflect differences in clinical practice because registries and registrars were not instructed to code tobacco-associated cancers differently than other cancers. This difference suggests that health care providers may document tobacco use history more frequently when they are diagnosing tobacco-associated cancers.

The percentages of never, current, and former tobacco users in our study were similar to national results from past studies.<sup>17–21</sup> In 2015, 15.1% of US adults reported being current cigarette smokers in the NHIS, whereas 17.6% of adults in this study were current cigarette smokers.<sup>7</sup> According to 2014 BRFSS data, the cigarette smoking prevalence ranged by state from 9.7% in Utah to 26.7% in West Virginia.<sup>18</sup> However, in addition to all cases having been diagnosed with cancer, the NPCR population is different from the NHIS and BRFSS populations in other respects, such as the age distribution and the number of states represented.<sup>22</sup> A telephone interview study of 20,891 patients with breast cancer during 1988–2008 reported a cigarette smoking prevalence of 20%,<sup>19</sup> whereas in our study, 17.6% of breast cancer cases were current cigarette users. Past literature has reported cigarette use prevalence above the national average in patients with tobacco-associated cancers such as lung and laryngeal cancers.<sup>20,21</sup> A 2012 study by Park et al<sup>20</sup> reported that 38.7% of patients with lung cancer had reported current smoking at diagnosis, and 90.2% were ever smokers; these figures are comparable to those for current (36.4%) and ever smokers (85.1%) among cases with tracheal, lung, or bronchus cancer in our study. The 2012 National Substance and Drug Use and Health survey data reported that 3.5% of US adults used smokeless tobacco in the past month; this is higher than the figure of 0.9% for smokeless tobacco current users reported in our study.<sup>23</sup>

Although survey-based tobacco use surveillance relies on the respondent’s recall of his or her tobacco use, studies show that the smoking status has been reliably self-reported.<sup>24</sup> Similarly, reviews of electronic medical records have indicated that current tobacco use is reliably recorded, although past tobacco use may be under-recorded.<sup>25,26</sup> Although the response rate of tobacco use questions in the BRFSS ranges from 25.1% to 60.1% and is dependent on the US state,<sup>18</sup> the rate of documentation of tobacco use in electronic medical records was 64.4% in a Canadian study of EHRs looking at 249,223 patients.<sup>27</sup> Because of the Meaningful Use Initiative in the United States, which encourages hospitals and clinics to record the smoking history in EHRs for at least 80% of patient encounters,<sup>28</sup> the recording of tobacco use in EHRs may be improving.<sup>29</sup> The development of standards for how tobacco use is most accurately and easily recorded and subsequently accessed for clinical or research use in EHRs may improve the ease of data collection for registrars and improve the percent known for tobacco use variables.



Data with a high percent known and high data validity have the potential to be used to better understand the relation between tobacco use and cancer outcomes such as progression, recurrence, secondary cancers, and mortality. Known tobacco use among cancer patients can also inform cancer control programs and the recruitment of patients for tobacco use cessation interventions.<sup>30</sup> Past studies have used cancer registries to recruit patients for outreach for cancer survivors and for breast cancer screening.<sup>31–33</sup> Tobacco use data from cancer registries could enable health professionals to identify patients who use tobacco at diagnosis or to track patients' tobacco use after diagnosis. However, ideal methods for contacting this patient population and recording the most recent smoking status will need further evaluation. A study using local cancer registries to identify patients for smoking cessation interventions found that 71% of 577 individuals listed as current cigarette smokers in the registry at cancer diagnosis were identified as current cigarette smokers during a follow-up abstraction of EHR records 6 to 24 months after their diagnosis.<sup>34</sup> In addition, standardized tobacco use information at diagnosis could be incorporated into clinical trials that use a follow-up evaluation to assess the impact of smoking on the prognosis.<sup>35</sup> Because the smoking status at and after diagnosis significantly affects the prognosis,<sup>2–4</sup> there is the potential to incorporate the smoking status into clinical decisions regarding tobacco cessation interventions or the management of survivorship care. In contrast, a low percentage of people were recorded as using other tobacco products; this may lead to challenges in identifying patients for intervention efforts and in delivering public health interventions and thus limit the usefulness of other tobacco use variables.

The collection of the tobacco use status through cancer registries depends on the accuracy of the health care providers recording tobacco use, the ease of the registrars in finding this information in the record, and the acceptability among state registries in collecting these data. In addition, because cancer cases are sometimes not abstracted until close to 6 months after the diagnosis and are not reported to the state until as late as 2 years after the diagnosis (depending on the state), some registries may have limited usefulness for enrolling patients in immediate smoking cessation interventions. Furthermore, increasing the data collection burden on cancer registrars may increase registry operating costs, especially because modifications to software would be needed to standardize tobacco use data collection.<sup>36</sup>

This study had several limitations. Although data from 10 central cancer registries were included in the CER study, the results may not be generalizable to other states. Second, despite registry quality standards and an overall high quality of data,<sup>37</sup> variables in the NPCR and associated registries may be subject to coding misclassification. The internal validity of tobacco use data and the degree of potential coding misclassification are not well documented in the literature and may vary by additional factors such as the reporting source, age, sex, and type of cancer diagnosis. The NPCR smoking variables did not assess the amount or duration of tobacco use (eg, pack-years), which is a valuable tool and is used by other organizations, including the Centers for Medicare and Medicaid Services (as part of lung cancer screening eligibility criteria) and the American Joint Committee on Cancer.<sup>38,39</sup> Pack-years were not included in the CER study because the level of detail needed to access this information accurately is not routinely and consistently available in the medical records from which abstractors collected the CER study data. In addition, it is known that EHRs underestimate pack-years or have missing pack-year data,<sup>40,41</sup> and in the context of a survey,

multiple questions are required for an accurate assessment.<sup>42</sup> Adding pack-years to CER tobacco use variables for registrar abstraction may require further data validation and standardization.<sup>40</sup> Finally, electronic cigarette (e-cigarette) use and nicotine gum use were not included in tobacco use variable data abstraction, although states were asked to note the use of these items in a text field for future reference. Although e-cigarette use has been increasing in prevalence,<sup>43</sup> the use of e-cigarettes presents a different toxicity profile than cigarette use<sup>44</sup> and may have further complicated the study's aim of collecting tobacco use data.

There may be several ways to improve the quality of tobacco use variables. First, because of the low percent known and low percentages of ever users reported for non-cigarette tobacco use variables, cancer registries could consider focusing on cigarette smoking only. The Meaningful Use Initiative asked only about cigarette smoking instead of all forms of tobacco use to focus on 1 quality objective that maximizes the public health impact.<sup>45</sup> Focusing on cigarette use may lead to higher variable acceptability and reliability. Second, the quality of tobacco use data collected by cancer registries is dependent on the documentation in the medical record. The education of providers as well as policies that incentivize documenting tobacco use may improve tobacco use documentation.<sup>46,47</sup> Third, standardization of tobacco use variables may be needed to improve variable quality and best address public health planning needs.<sup>35</sup> Finally, further data validation will be needed. The wide variation in results by state may suggest a need for a further evaluation of methods and quality registry reporting. Future studies could internally validate tobacco use variables by comparing registry data with a record review or patient interview gold standard.

In conclusion, tobacco use data collected through cancer registries could strengthen public health efforts in reducing the burden of tobacco-associated cancers. Continued partnership between the cancer registry community and initiatives such as the Meaningful Use Initiative that better define EHR variables will increase variable quality to the benefit of health care providers, patients, public health planners, and researchers.

## Acknowledgments

We acknowledge the project investigators at the participating central cancer registries as well as other organizations and the individuals, including the registrars, who supported the collection of the data to enhance the National Program of Cancer Registries for Comparative Effectiveness Research: Judy Brockhouse (Alaska Cancer Registry), Dee W. West (Cancer Registry of Greater California), Randi K. Rycroft (Colorado Central Cancer Registry), Christopher J. Johnson (Cancer Data Registry of Idaho), Monique N. Hernandez (Florida Cancer Data System), Christie R. Ehemann and Timothy S. Styles (Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention), Kevin B. Zhang (ICF International), Vivien Chen and Xiao-Cheng Wu (Louisiana Tumor Registry and Epidemiology Program), David Rousseau (Rhode Island Cancer Registry), Maria O. Celaya (New Hampshire State Cancer Registry), Jennifer M. Wike (Centers for Disease Control and Prevention–National Program of Cancer Registries contractor DB Consulting), Melissa Pearson (North Carolina Cancer Registry), and Anne M. Hakenewerth (Texas Cancer Registry).

### FUNDING SUPPORT

This work was supported in part under Centers for Disease Control and Prevention cooperative agreements of the National Program of Cancer Registries (U58/DP000792) in conjunction with the participating states and a Centers for Disease Control and Prevention comparative effectiveness research contract to ICF (200-2008-27957).

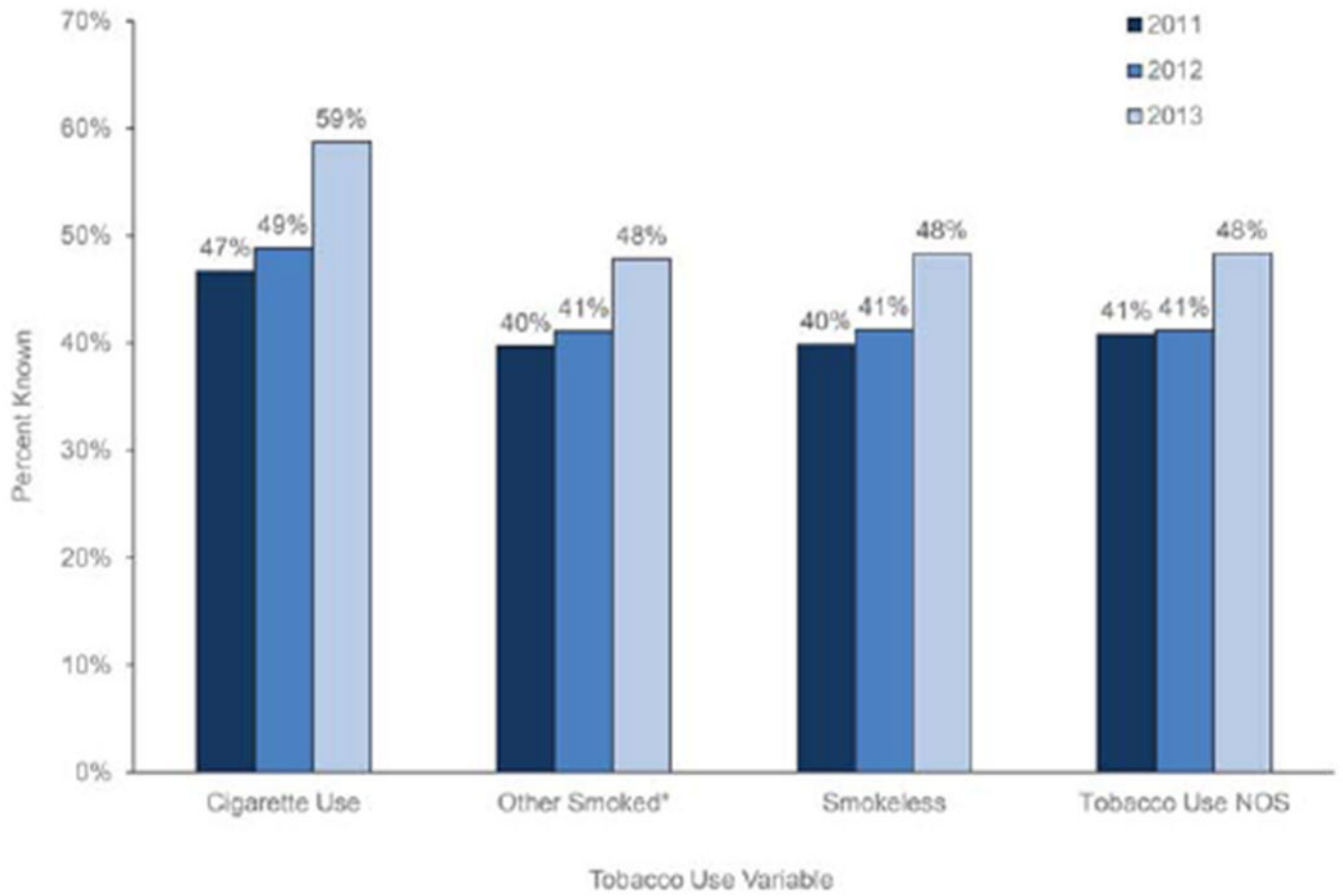


## REFERENCES

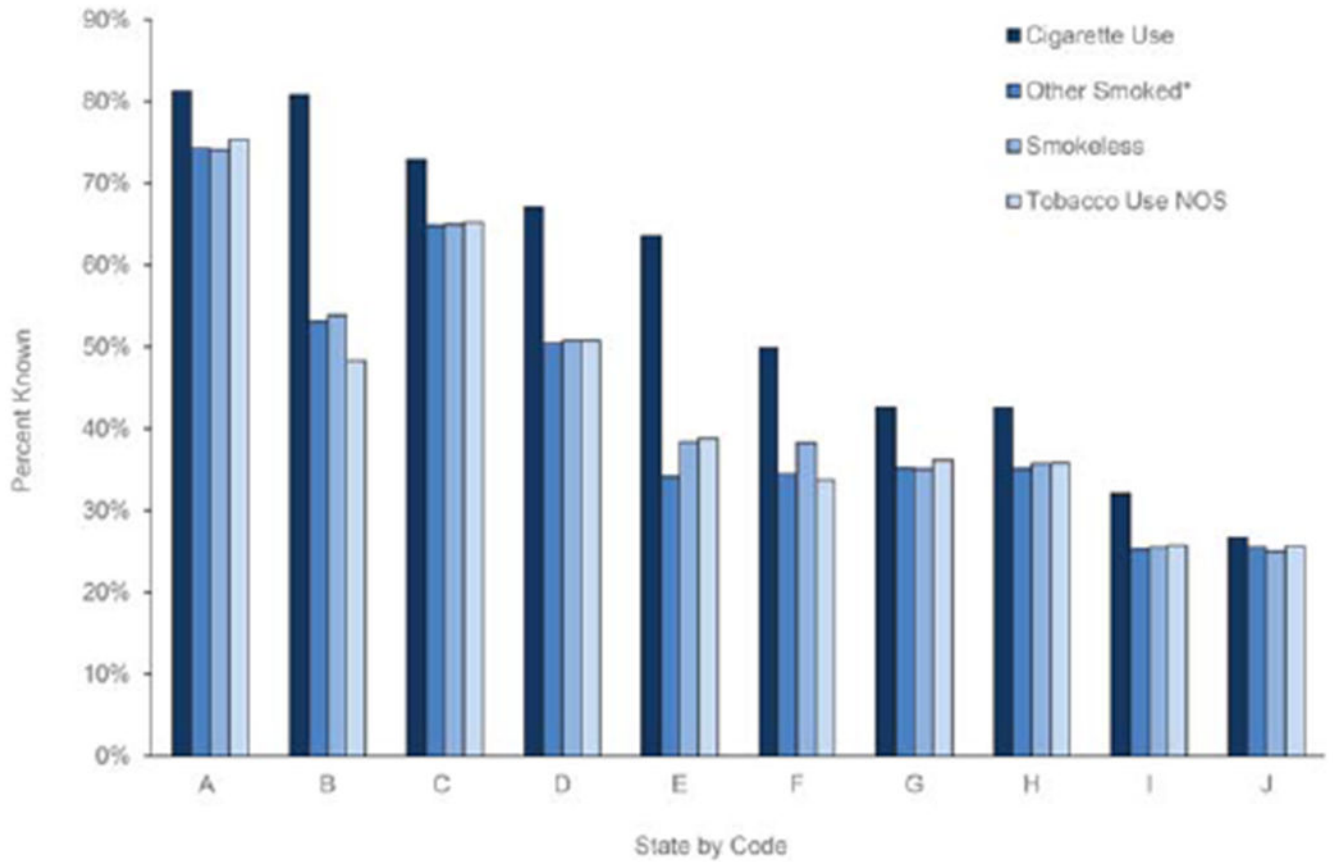
1. Siegel RL, Jacobs EJ, Newton CC, et al. Deaths due to cigarette smoking for 12 smoking-related cancers in the United States. *JAMA Intern Med.* 2015;175:1574–1576. [PubMed: 26076120]
2. Parsons A, Daley A, Begh R, Aveyard P. Influence of smoking cessation after diagnosis of early stage lung cancer on prognosis: systematic review of observational studies with meta-analysis. *BMJ.* 2010; 340:b5569. [PubMed: 20093278]
3. Underwood JM, Townsend JS, Tai E, White A, Davis SP, Fairley TL. Persistent cigarette smoking and other tobacco use after a tobacco-related cancer diagnosis. *J Cancer Surviv.* 2012;6:333–344. [PubMed: 22706885]
4. Ferketich AK, Niland JC, Mamet R, et al. Smoking status and survival in the National Comprehensive Cancer Network non-small cell lung cancer cohort. *Cancer.* 2013;119:847–853. [PubMed: 23023590]
5. Wyszynski A, Tanyos SA, Rees JR, et al. Body mass and smoking are modifiable risk factors for recurrent bladder cancer. *Cancer.* 2014;120:408–414. [PubMed: 24122218]
6. Sharp L, McDevitt J, Brown C, Carsin AE, Comber H. 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.* 2017;123:2543–2550. [PubMed: 28297071]
7. Delnevo CD, Bauer UE. Monitoring the tobacco use epidemic III: the host: data sources and methodological challenges. *Prev Med.* 2009;48(1 suppl):S16–S23. [PubMed: 18851990]
8. Warren GW, Kasza KA, Reid ME, Cummings KM, Marshall JR. Smoking at diagnosis and survival in cancer patients. *Int J Cancer.* 2013;132:401–410. [PubMed: 22539012]
9. Chen VW, Ehemann CR, Johnson CJ, et al. Enhancing cancer registry data for comparative effectiveness research (CER) project: overview and methodology. *J Registry Manag.* 2014;41:103–112. [PubMed: 25419602]
10. World Health Organization. *International Classification of Diseases for Oncology.* 3rd ed, 1st rev Geneva, Switzerland: World Health Organization; 2013.
11. 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; 2014.
12. Henley SJ, Thomas CC, Sharapova SR, et al. Vital signs: disparities in tobacco-related cancer incidence and mortality—United States, 2004–2013. *MMWR Morb Mortal Wkly Rep.* 2016;65:1212–1218. [PubMed: 27832048]
13. Surveillance, Epidemiology, and End Results. Site recode. <https://seer.cancer.gov/siterecode/> Accessed July 17, 2017.
14. Centers for Disease Control and Prevention. National Program of Cancer Registries comparative effectiveness research data set: 2011 CER restricted access data dictionary. [https://www.cdc.gov/rdc/data/b1/CDC\\_DCPC\\_CancerSurveillanceBranch\\_CER.pdf](https://www.cdc.gov/rdc/data/b1/CDC_DCPC_CancerSurveillanceBranch_CER.pdf) Accessed January 19, 2018.
15. Wilson RJ, O’Neil ME, Ntekop E, Zhang K, Ren Y. Coding completeness and quality of relative survival-related variables in the National Program of Cancer Registries cancer surveillance system, 1995–2008. *J Registry Manag.* 2014;41:65–71. [PubMed: 25153011]
16. Knowlton R, Gershman S, Solis A, Das B. An assessment of the reliability of race, Hispanic ethnicity, birthplace, and tobacco history data in the Massachusetts cancer registry, 2005–2009. *J Registry Manag.* 2014;41:146–150. [PubMed: 25419609]
17. Phillips E, Wang TW, Husten CG, et al. Tobacco product use among adults—United States, 2015. *MMWR Morb Mortal Wkly Rep.* 2017;66:1209–1215. [PubMed: 29121001]
18. Nguyen KH, Marshall L, Brown S, Neff L. State-specific prevalence of current cigarette smoking and smokeless tobacco use among adults—United States, 2014. *MMWR Morb Mortal Wkly Rep.* 2016; 65:1045–1051. [PubMed: 27711031]
19. 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.* 2016;34:1315–1322. [PubMed: 26811527]

20. Park ER, Japuntich SJ, Rigotti NA, et al. A snapshot of smokers after lung and colorectal cancer diagnosis. *Cancer*. 2012;118:3153–3164. [PubMed: 22271645]
21. Choi SH, Terrell JE, Bradford CR, et al. Does quitting smoking make a difference among newly diagnosed head and neck cancer patients? *Nicotine Tob Res*. 2016;18:2216–2224. [PubMed: 27613928]
22. Johnson CJ, Murphy CF, Rich M, Cariou C, Mastalski K. Health status and behaviors among cancer survivors in Idaho, 2014 [http://www.ccaidaho.org/Health%20Status%20and%20Behaviors%20among%20Cancer%20Survivors%20in%20Idaho%202014\\_FINAL\\_11212016.pdf](http://www.ccaidaho.org/Health%20Status%20and%20Behaviors%20among%20Cancer%20Survivors%20in%20Idaho%202014_FINAL_11212016.pdf) Accessed January 19, 2018.
23. Substance Abuse and Mental Health Services Administration. The NSDUH report: trends in smokeless tobacco use and initiation: 2002 to 2012. [https://www.samhsa.gov/data/sites/default/files/189\\_NSDUH\\_Trends\\_Smokeless\\_Tobacco/NSDUH-SR189-SmokelessTob-2014.htm](https://www.samhsa.gov/data/sites/default/files/189_NSDUH_Trends_Smokeless_Tobacco/NSDUH-SR189-SmokelessTob-2014.htm) Accessed December 6, 2017.
24. Brigham J, Lessov-Schlaggar CN, Javitz HS, McElroy M, Krasnow R, Swan GE. Reliability of adult retrospective recall of lifetime tobacco use. *Nicotine Tob Res*. 2008;10:287–299. [PubMed: 18236293]
25. Booth HP, Prevost AT, Gulliford MC. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. *Pharmacoepidemiol Drug Saf*. 2013;22:1357–1361. [PubMed: 24243711]
26. Lewis JD, Brensinger C. Agreement between GPRD smoking data: a survey of general practitioners and a population-based survey. *Pharmacoepidemiol Drug Saf*. 2004;13:437–441. [PubMed: 15269927]
27. Greiver M, Aliarzadeh B, Meaney C, et al. Are we asking patients if they smoke?: missing information on tobacco use in Canadian electronic medical records. *Am J Prev Med*. 2015;49:264–268. [PubMed: 25997907]
28. Centers for Medicare and Medicaid Services. Stage 2 eligible professional meaningful use core measures. [https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/Stage2\\_EPCore\\_5\\_RecordSmokingStatus.pdf](https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/Stage2_EPCore_5_RecordSmokingStatus.pdf) Accessed January 7, 2017.
29. Levine DM, Healey MJ, Wright A, Bates DW, Linder JA, Samal L. Changes in the quality of care during progress from stage 1 to stage 2 of meaningful use. *J Am Med Inform Assoc*. 2017;24:394–397. [PubMed: 27567000]
30. Hartman AM, Thun MJ, Ballard-Barbash R. Linking tobacco control policies and practices to early cancer endpoints: surveillance as an agent for change. *Cancer Epidemiol Biomarkers Prev*. 2008;17:2215–2219. [PubMed: 18768485]
31. Carpentier MY, Tiro JA, Savas LS, et al. Are cancer registries a viable tool for cancer survivor outreach? A feasibility study. *J Cancer Surviv*. 2013;7:155–163. [PubMed: 23247719]
32. Katapodi MC, Northouse LL, Schafenacker AM, et al. Using a state cancer registry to recruit young breast cancer survivors and high-risk relatives: protocol of a randomized trial testing the efficacy of a targeted versus a tailored intervention to increase breast cancer screening. *BMC Cancer*. 2013;13:97. [PubMed: 23448100]
33. Arora NK, Hamilton AS, Potosky AL, et al. Population-based survivorship research using cancer registries: a study of non-Hodgkin's lymphoma survivors. *J Cancer Surviv*. 2007;1:49–63. [PubMed: 18648945]
34. Krebs P, Rogers E, Wong H, Ostroff JS, Henley SJ. Assessing the accuracy of registry-based tobacco use status and utility for patient recruitment into tobacco trials. <https://20tqtx36s11a18rvn82wcmpn-wpengine.netdna-ssl.com/wp-content/uploads/2017/06/NAACCR-2017-Abstract-Final-Program.pdf> Accessed December 7, 2017.
35. Land SR, Toll BA, Moinpour CM, et al. Research priorities, measures, and recommendations for assessment of tobacco use in clinical cancer research. *Clin Cancer Res*. 2016;22:1907–1913. [PubMed: 26888828]
36. Tangka FK, Subramanian S, Beebe MC, et al. Cost of operating central cancer registries and factors that affect cost: findings from an economic evaluation of Centers for Disease Control and Prevention National Program of Cancer Registries. *J Public Health Manag Pract*. 2016;22:452–460. [PubMed: 26642226]

37. Thoburn KK, German RR, Lewis M, Nichols PJ, Ahmed F, Jackson-Thompson J. Case completeness and data accuracy in the Centers for Disease Control and Prevention's National Program of Cancer Registries. *Cancer*. 2007;109:1607–1616. [PubMed: 17343277]
38. Fucito LM, Czabafy S, Hendricks PS, Kotsen C, Richardson D, Toll BA. Pairing smoking-cessation services with lung cancer screening: a clinical guideline from the Association for the Treatment of Tobacco Use and Dependence and the Society for Research on Nicotine and Tobacco. *Cancer*. 2016;122:1150–1159. [PubMed: 26916412]
39. Ridge JA, Lydiatt WM, Snehal GP, et al. Lip and oral cavity In: Amin MB, Edge SB, Greene FL, et al., eds. *AJCC Cancer Staging Manual*. 8th ed. Cham, Switzerland: Springer; 2016:79–94.
40. Modin HE, Fathi JT, Gilbert CR, et al. Pack-year cigarette smoking history for determination of lung cancer screening eligibility. Comparison of the electronic medical record versus a shared decision-making conversation. *Ann Am Thorac Soc*. 2017;14:1320–1325. [PubMed: 28406708]
41. Kinsinger LS, Anderson C, Kim J, et al. Implementation of lung cancer screening in the Veterans Health Administration. *JAMA Intern Med*. 2017;177:399–406. [PubMed: 28135352]
42. Centers for Disease Control and Prevention. 2017 Behavioral Risk Factor Surveillance System questionnaire. [https://www.cdc.gov/brfss/questionnaires/pdf-ques/2017\\_BRFSS\\_Pub\\_Ques\\_508\\_tagged.pdf](https://www.cdc.gov/brfss/questionnaires/pdf-ques/2017_BRFSS_Pub_Ques_508_tagged.pdf) Accessed January 19, 2018.
43. King BA, Patel R, Nguyen KH, Dube SR. Trends in awareness and use of electronic cigarettes among US adults, 2010-2013. *Nicotine Tob Res*. 2015;17:219–227. [PubMed: 25239961]
44. Goniewicz ML, Knysak J, Gawron M, et al. Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. *Tob Control*. 2014;23:133–139. [PubMed: 23467656]
45. Meaningful Use Workgroup. Draft transcript. [https://www.healthit.gov/archive/archive\\_files/FACA%20Hearings/2011/2011-04-05%20Meaningful%20Use/2011-04-05\\_policy\\_mu\\_transcript\\_draft.pdf](https://www.healthit.gov/archive/archive_files/FACA%20Hearings/2011/2011-04-05%20Meaningful%20Use/2011-04-05_policy_mu_transcript_draft.pdf) Accessed February 9, 2017.
46. Kruse GR, Chang Y, Kelley JH, Linder JA, Einbinder JS, Rigotti NA. Healthcare system effects of pay-for-performance for smoking status documentation. *Am J Manag Care*. 2013;19: 554–561. [PubMed: 23919419]
47. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. *Am J Prev Med*. 2008;35: 158–176. [PubMed: 18617085]



**Figure 1.** Percent known for tobacco use variables among cancer cases in 10 central cancer registries, 2011-2013. The counts for the tobacco use cigarette variable excluded 3 cases with blank values. The tobacco use other smoked variable excluded 3 blank values, the tobacco use smokeless variable excluded 3 blank values, and the tobacco use NOS variable excluded 4 blank values. \*Smoked tobacco products other than cigarettes (eg, pipes, cigars, and kreteks). NOS indicates not otherwise specified.



**Figure 2.** Percent known for tobacco use variables recorded among cancer cases (n = 1,646,505) in 10 central cancer registries by coded state, 2011-2013. The counts of the tobacco use cigarette variable excluded 3 cases with blank values. The tobacco use other smoked variable excluded 3 blank values, the tobacco use smokeless variable excluded 3 blank values, and the tobacco use NOS variable excluded 4 blank values. \*Smoked tobacco products other than cigarettes (eg, pipes, cigars, and kreteks). NOS indicates not otherwise specified.

**TABLE 1.**  
 Characteristics and Percent Known of Cancer Cases in 10 Central Cancer Registries, 2011-2013

Characteristic	Percent Known of Tobacco Use Variables <sup>a</sup>			
	Count, No. <sup>b</sup>	Cigarettes	Other Smoked <sup>c</sup>	Smokeless NOS
Total	1,646,505	51.4	42.9	43.1
Sex				
Male	812,096	49.6	40.5	40.7
Female	834,409	53.1	45.1	45.4
Diagnosis age (median, 66 y)				
0-29 y	41,856	50.3	45.2	45.4
30-39 y	54,164	50.7	43.4	43.6
40-49 y	140,928	52.2	44.4	44.7
50-59 y	316,138	52.3	43.5	43.8
60-69 y	453,141	51.8	42.8	43.0
70-79 y	377,652	52.2	43.1	43.4
80 y	262,626	48.4	40.5	40.7
Race				
White	1,365,514	52.2	43.4	43.7
Black	166,620	55.2	47.0	47.1
Asian/Pacific Islander	73,157	38.9	32.8	32.9
Other	41,214	30.9	25.5	25.7
Ethnicity <sup>d</sup>				
Non-Hispanic	1,410,097	51.7	42.9	43.2
Hispanic	235,773	49.5	42.5	42.4
Diagnosis type				
Oral cavity and pharynx	40,994	55.8	46.0	46.4
Digestive system	277,305	54.6	45.4	45.6
Respiratory system	205,806	58.1	45.2	45.4
Skin, excluding basal and squamous	128,608	26.9	22.6	23.1
Breast	268,728	56.0	48.4	48.5



Characteristic	Percent Known of Tobacco Use Variables <sup>a</sup>				
	Count, No. <sup>b</sup>	Cigarettes	Other Smoked <sup>c</sup>	Smokeless	NOS
Female genital system	87,334	54.0	46.7	47.0	47.2
Male genital system	191,235	47.0	39.0	39.1	39.9
Urinary system	120,881	51.8	42.3	42.8	43.0
Brain and other nervous system	60,179	52.8	45.1	45.4	45.1
Endocrine system	58,090	49.8	42.3	42.8	42.3
Blood cancer	127,991	52.4	44.4	44.5	44.7
Other	79,354	50.4	42.8	42.9	43.2
Behavior					
Malignant	1,475,270	52.6	43.8	44.0	44.4
In situ	118,758	36.4	31.1	31.5	31.3
Borderline malignancy	4616	49.7	43.1	43.4	43.0
Benign	47,861	51.4	43.9	44.4	43.9
Reporting source					
Hospital	1,462,351	54.6	45.6	45.9	46.1
Radiation or oncology center	30,959	72.8	61.1	59.8	64.8
Laboratory only	27,136	9.4	7.7	7.7	7.7
Physician office/private practitioner	69,734	9.0	7.4	7.5	7.6
Nursing/convalescent home/hospice	1020	2.6	1.6	1.6	1.9
Autopsy only	984	16.2	11.9	11.7	12.4
Death certificate only	23,851	0.2	0.2	0.2	0.4
Other hospital center	30,470	51.1	41.2	41.2	41.1

Abbreviation: NOS, not otherwise specified.

<sup>a</sup>The percent known was calculated by the division of the total known values of each tobacco use variable by the number of cases as represented in the count column. The chi-squared *P* value within each characteristic and for each tobacco use variable was <.0001. The counts for the tobacco use cigarette variable excluded 3 cases with blank values. The tobacco use other smoked variable excluded 3 blank values, the tobacco use smokeless variable excluded 3 blank values, and the tobacco use NOS variable excluded 4 blank values.

<sup>b</sup>Cases in this column served as the denominator for calculating the row percentages for the percent known.

<sup>c</sup>Smoked tobacco products other than cigarettes (eg, pipes, cigars, and kreteks).

<sup>d</sup>Excluded 635 cases of unspecified ethnicity.

**TABLE 2.**  
Tobacco Use Variables Recorded in 10 Central Cancer Registries, 2011–2013, Excluding Cases With an Unknown Status

Response	Tobacco Variable Category, No. of Cases (%) <sup>a</sup>			NOS (n = 714,844)
	Cigarettes (n = 846,212)	Other Smoked (n = 705,639) <sup>b</sup>	Smokeless (n = 709,635)	
Never user	432,113 (51.1)	686,071 (97.2)	696,404 (98.1)	666,995 (93.3)
Current user	149,063 (17.6)	9980 (1.4)	6493 (0.9)	17,373 (2.4)
Former user	265,036 (31.3)	9588 (1.4)	6738 (1.0)	30,476 (4.3)
Quit within 1 y from diagnosis	22,574 (2.7)	1190 (0.2)	1007 (0.1)	2729 (0.4)
Quit more than 1 y before diagnosis	199,764 (23.6)	5984 (0.8)	3879 (0.6)	22,509 (3.1)
Unknown time since quitting	42,698 (5.0)	2414 (0.3)	1852 (0.3)	5238 (0.7)

Abbreviation: NOS, not otherwise specified.

<sup>a</sup>The tobacco use cigarette variable excluded 800,290 unknowns and 3 cases with blank values, the tobacco use other smoked variable excluded 940,863 unknowns and 3 blanks, the tobacco use smokeless variable excluded 936,867 unknowns and 3 blanks, and the tobacco use NOS variable excluded 931,657 unknowns and 4 blanks.

<sup>b</sup>Smoked tobacco products other than cigarettes (eg, pipes, cigars, and kreteks).

TABLE 3.

Cigarette Use Recorded in 10 Central Cancer Registries by Tobacco-Associated Cancer Type, 2011-2013

Cancer Type	Total Cases, No.	Total Known, No. <sup>a</sup>	Percent Known <sup>b</sup>	Ever User, % <sup>c</sup>	Current User, %	Former User, %	Never User, %
All cancers	1,646,502	846,212	51.4	48.9	17.6	31.3	51.1
Non-tobacco-associated cancers	1,017,984	498,409	49.0	39.0	12.4	26.6	61.0
Tobacco-associated cancers	628,518	347,803	55.3	63.2	25.1	38.1	36.8
Larynx	11,849	7067	59.6	86.2	44.6	41.6	13.8
Trachea, lung, and bronchus	188,386	109,438	58.1	85.1	36.4	48.8	14.9
Esophagus	14,736	8282	56.2	70.9	26.2	44.7	29.1
Lip, oral cavity, and pharynx	40,945	22,855	55.8	65.7	29.1	36.6	34.3
Urinary bladder	62,352	32,362	51.9	63.8	21.0	42.8	36.2
Liver	28,841	14,154	49.1	62.8	27.6	35.1	37.2
Stomach	23,405	12,187	52.1	50.4	16.5	33.9	49.6
Pancreas	41,363	21,994	53.2	49.2	17.0	32.3	50.8
Kidney and renal pelvis	54,928	28,318	51.6	48.8	17.5	31.2	51.2
Colon and rectum	135,236	76,438	56.5	43.7	15.1	28.6	56.3
Acute myeloid leukemia	13,321	7448	55.9	41.6	11.8	29.8	58.4
Cervix uteri	13,156	7260	55.2	41.5	24.2	17.3	58.5

<sup>a</sup>Cases in this column served as the denominator for calculating the row percentages of ever, current, former, and never cigarette users.<sup>b</sup>The percent known was calculated by the division of the total known cases by the total cases.<sup>c</sup>Ever users represent the sum of the current and former user columns.