Tobacco-Use Prevalence in Special Populations

Taking Advantage of Electronic Medical Records

Leif I. Solberg, MD, Thom J. Flottemesch, PhD, Steven S. Foldes, PhD, Beth A. Molitor, BS, Patricia F. Walker, MD, DTM&H, A. Lauren Crain, PhD

Background: It is difficult and expensive to use surveys to obtain the repeatable information that is needed to understand and monitor tobacco prevalence rates and to evaluate cessation interventions among various subgroups of the population. Therefore, the electronic medical record database of a large medical group in Minnesota was used to demonstrate the potential value of that approach to accomplish those goals.

Methods: The relevant variables for all medical group patients aged 18 and over were extracted from the record from a 1-year period. Rates of smoking prevalence were computed for the entire population as well as for those with various characteristics and combinations of characteristics of interest to tobacco-cessation advocates. These prevalence rates were also adjusted to control for the other characteristics in the analysis.

Results: From March 2006 to February 2007, there were 183,982 unique patients with at least one office visit with a clinician, and a record of their tobacco-use status (90%). Overall, 19.7% with recorded status were tobacco users during this year, as were 24.2% of those aged 18–24 years, 16.0% of pregnant women, 34.3% of those on Medicaid, 40.0% of American Indians, 9.5% of Asians, and 8.5% of those whose preferred language was other than English. Combining characteristics allowed greater understanding of those differences.

Conclusions: Although there are limitations in these data, the level of detail available for this large population and the ease of repeat analysis should greatly facilitate targeted interventions and evaluation of the impact.

(Am J Prev Med 2008;35(6S):S501–S507)

Introduction

There is relatively good information about the national prevalence of tobacco use by various subgroups of the adult population. However, comparable information is largely unavailable at the state level where tobacco-control interventions are most likely and most effective. Neither the National Health Interview Survey (NHIS), the National Ambulatory Medical Care Survey (NAMCS), nor the Behavioral Risk Factor Surveillance Survey (BRFSS) provide such information.[1–3] Even if a larger and more representative sample were available, measuring rates by surveys would still be prohibitively costly and doing so frequently is beyond current possibility. Thus far, tobacco-use prevalence data have all been based on such survey methods.

The absence of accurate and frequently repeatable estimates of tobacco-use prevalence rates among specific populations of a state or region makes it difficult to develop specific tobacco-control interventions, much less to monitor changes over time as a measure of effectiveness. Thus, there is a need for such information from some other source.

Therefore, we proposed to use the data in a large medical group's electronic medical record (EMR) system for this purpose. That task became possible in a medical group in the Minneapolis/St. Paul MN metropolitan area because it had:

1. implemented a relatively consistent collection of electronically accessible data about the race, ethnicity, and language preference of most patients;
2. accessible data for nearly all patients on their age, gender, pregnancy, health insurance status, and tobacco use;
3. a large patient population with a distribution in terms of race, ethnicity, age, and income status similar to that of the urban metropolitan counties it serves as well as that of the state as a whole; and

From HealthPartners Research Foundation and Medical Group (Solberg, Flottemesch, Molitor, Walker, Crain), Minneapolis, Minnesota; and Blue Cross and Blue Shield of Minnesota (Foldes), St. Paul, Minnesota

Address correspondence and reprint requests to: Leif I. Solberg, MD, HealthPartners Research Foundation, P.O. Box 1524, MS#21111R, Minneapolis MN 55440-1524. E-mail: leif.i.solberg@healthpartners.com.

doi:10.1016/j.amepre.2008.08.033

4. created a shadow system of its complete medical records that was updated daily and was available for data analysis without disturbing use of the record system for patient care.

It was decided to use this record system to measure tobacco-use prevalence among young adults (aged 18–24 years); pregnant women; people on Medicaid as a surrogate for low income; and people of various races, ethnicities, and language preferences. This paper reports the findings on various tobacco-prevalence rates, as well as more detailed analyses to explore the potential to adjust these rates for multiple overlaps among the populations studied.

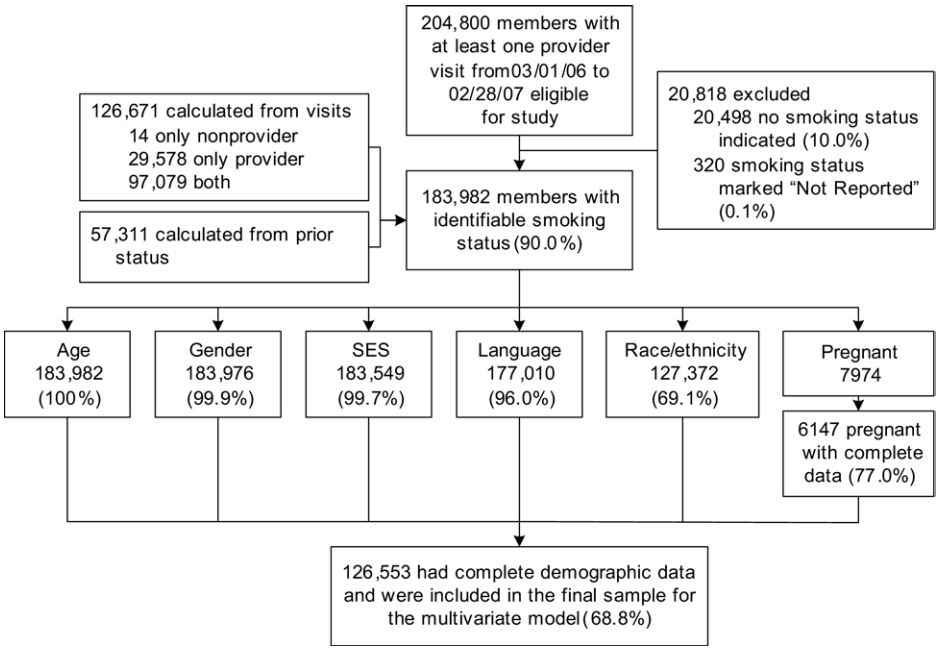

Figure 1. Data flow diagram

Methods

HealthPartners Medical Group (HPMG) is a 650-physician multi-specialty care delivery organization that is affiliated with a health plan with 750,000 members. Although 60% of its patients have HealthPartners health insurance, the other 40% have a wide variety of other insurance, including Medicaid or self-pay. The overall patient group in 2007 included 18% nonwhite and 11% on various Medicaid programs, compared to 14% nonwhite and 10% on Medicaid for the entire state and 24% and 9% for the two main urban counties in which most of the medical group care sites are located (see www.census.gov).

Beginning in 2000, HPMG implemented a paperless electronic medical record (Epic®) in all its clinics. In 2002, identification of each patient's race/ethnicity and preferred language were added to the tasks of registering and rooming patients as part of a major initiative to identify and address disparities. Because of the difficulty and time required for staff to ask patients separately about race and ethnicity, Hispanic/Latino was simply listed as one of the available alternative categories.

Data

All data came from the shadow server of HPMG's Epic® EMR system, where data are stored at either the patient or the visit level. For visit-level data, a new record for each data element is created at each visit, so it is possible to track changes in data elements over time. In contrast, patient-level data are overwritten at each visit, so only the most recent update is available. Prior to February 2006, tobacco status was stored at the patient level. Since that did not allow changes in individual tobacco-use status to be determined, the 12-month time frame of March 2006 to the end of February 2007 was used for this study.

Figure 1 provides a summary picture of the number and percent of people included in this analysis. Data were collected for all individuals aged ≥18 as of January 1, 2006, with at least one provider visit from March 2006 to the end of February 2007. This identified 204,800 individuals with 690,054 provider visits out of a total population of 237,822 (86%) active patients. The provider limitation for this analysis was because collecting demographic information including tobacco status is not as consistent a part of the workflow of nonprovider visits. However, these 204,800 individuals also had 134,388 nonprovider visits with a recorded tobacco status, so all visits were used when determining each person's tobacco-use status.

In addition to tobacco-use status, the following visit-level variables were also extracted from the EMR during this time frame: date of visit; department; clinic; insurance type (commercial, Medicare, Medicaid, Medicare + Medicaid, Minnesota Care [Minnesota Care is a unique state program to provide health coverage to the uninsured], or other); payer; and whether the visit was for prenatal care. These data elements were used to construct patient-level variables for the analysis. Individuals were defined as low income if their insurance type was Medicaid, Medicare + Medicaid, or Minnesota Care across all visits. The dates of the first and last identified prenatal visit as well as the estimated date of confinement were used to define a 9-month pregnancy window.[4] For a pregnancy comparison group, a 1:1 cohort of nonpregnant women was matched by age, ethnicity, and preferred language. The following patient-level variables were collected: age, gender, race/ethnicity, and preferred language, as well as each individual's patient-level tobacco status as February 2006.

Determining Tobacco-Use Status

Since February 2006, tobacco-use status has been categorized as *never, yes, quit, passive* (exposed to the smoke of others), or

not asked for any form of tobacco (not just cigarette) use. Since the passive category did not have a clear interpretation and since there are no clear guidelines regarding its use, it was not included in these analyses.

Since February 2006 when tobacco use began to be stored as a visit-level variable, the last available tobacco status is automatically displayed at the start of each visit, but nothing is recorded for a visit unless someone actively inputs information at that visit. Unfortunately, of the 690,054 identified provider visits during the study time frame, 557,030 (81%) did not have tobacco status recorded. Most status-unrecorded visits (76.5%) corresponded to patients with a previous status of never using tobacco or of having quit. To address this problem, nonprovider visits with a recorded tobacco status were also reviewed. This added an additional 134,388 visits for a total of 267,412 visits with a recorded tobacco status, allowing smoking status categorization of 90% of adult patients with provider visits. Although an algorithm was developed to identify changes in tobacco status during the study period, there were so few of these and so much missing data that could have changed that identification, so those in the categories that correspond to a change in status were all collapsed into the category of current user for measuring prevalence of tobacco use.

Statistical Analysis

All statistical tests were conducted at the 5% level and performed using the SAS system. Due to the large sample size, rules of clinical rather than statistical significance were used. Effects were retained if associated with a >10% difference in the likelihood of being a former or current tobacco user.

Bivariate analyses identified which variables correlated with the recording of tobacco status. Variables associated with the reporting of tobacco status were identified by either contingency table or Fisher's exact test. The variables screened with contingency tables were: age (by 10-year category); gender; insurance; race/ethnicity; and pregnancy. A Cochran–Armitage test for trend assessed whether the reporting of tobacco status increased linearly with age. Fisher's exact tests compared minority racial/ethnic groups to whites, non–English to English speakers, men to women, and those with government sponsored insurance (Medicaid, Minnesota Care, or Medicare) to those with a commercial product.

Two multivariate, multinomial logistic regressions modeled the likelihood of tobacco use. Both models compared the categories of current user and former user to never user. Appropriate adjustments (truncating and/or centering on mean or median) were made to continuous variables as needed. The first model used all 126,553 people. The second used a matched cohort of 12,294 where each pregnant woman (n=6147) was matched to a nonpregnant woman based on age, race/ethnicity, and primary language. The models yielded predicted rates for each group, which were compared to that of a common reference group.

Results

Figure 1 summarizes the studied population and the availability of demographic data within that population. Of the 183,982 people identified as having known tobacco status and at least one provider visit during the study time-frame, 57,311 (31%) had their tobacco status contributed from visits prior to the study period. The demographic category with the lowest level of reporting (69.1%; n=127,372) was race/ethnicity. Overall, 126,553 people (70% or those with a reported smoking status) had complete data across all of the demographic categories.

Table 1 lists the frequencies for each of the variables described above for the 183,982 adult patients with tobacco-use status known, along with the sample sizes and tobacco status reporting rates. The smallest category consisted of those having both Medicare and Medicaid (dual coverage), with 877 people (data not shown in table). Even this relatively small sample still provided 80% power to detect a difference in tobacco-use prevalence of at least 5%. Although there was a significant difference in tobacco-use status reporting frequency among several populations due to large sample sizes, none of these differences rose to the level of clinical significance except for those with missing data. All identified groups had over 88% of members with a reported tobacco status. The group with the lowest rate of reporting was aged 18–24 (88.4%) and had commercial health insurance (89.1%). Underreporting of tobacco status appears to be correlated with the absence of other data such as insurance information (84.4%), ethnicity (80%), and preferred language (73.6%).

Table 2 reports tobacco-user status for all of the individuals with complete information by each patient-level factor. The right column contains results from bivariate analyses comparing current tobacco status to the listed factor. As shown, tobacco use decreases with age (p<0.001); decreases with pregnancy (p<0.001); is least common among Asians (p<0.001); and is much less common among those with a non–English language preference (p<0.001). In contrast, tobacco use is most common among those aged 18–24 years; those on Medicaid (p<0.001); American Indians (p<0.001); and African Americans (p<0.001). Rates of past tobacco use and successful cessation also differ significantly across groups. Women (p<0.001); Medicare enrollees (p<0.001); whites (p=0.09); American Indians (p<0.001); English speakers (p<0.001); and those belonging to older age groups are more likely to be former smokers.

A multinomial logistic regression was developed to more accurately compare expected tobacco prevalence among specific populations. The following factors were screened for incorporation into the final multivariate model: age, gender, insurance status, number of primary care visits, number of prescription drug orders, age as of March 1, 2007, ethnicity, and pregnancy status. Factors found to correlate positively with current tobacco use were gender, having Medicaid insurance, having English as a preferred language, and being American Indian. Factors found to correlate negatively with current tobacco use were being pregnant, older, Asian, or black. However, these main effects were mitigated by several inter-

actions. For instance, non–English-speaking black men were approximately half (OR 0.59; $p=<0.001$) as likely to be current tobacco users as English-speaking black men.

The factors positively associated with former tobacco use were being male ($p<0.001$); on Medicaid ($p<0.001$); English-speaking ($p<0.001$); pregnant ($p<0.001$); and American Indian ($p<0.001$). Factors negatively associated with former tobacco use were being Asian ($p<0.001$) or black ($p<0.001$). Once again, a series of interactions confounded the main effects. For instance, English-speaking blacks were 1.5 times as likely to be former tobacco users as non–English-speaking blacks, while those aged 18–24 years were much less likely to be former users than those aged 45–54 years (OR=0.37; $p<0.001$).

Table 3 shows the current- and former-user rates for each characteristic in the sample of people with complete information about each characteristic. This sample of 126,553 people has 57,429 fewer people (31.2%) than the overall population, so this table shows how their tobacco-use rates compare to those in the overall population. While there are some small differences in rates except for the race/ethnicity sample and pregnant women, these differences are quite small.

Table 1. Number and proportion of patients with various characteristics (N=183,982)

Characteristics	*n* with reported tobacco status	% of final sample	Tobacco status reporting rate	*p* value
Age (years)[a]				
18–24	20,029	10.9	88.4	<0.001
25–34	29,875	16.2	88.7	
35–44	35,297	19.2	88.0	
45–54	40,493	22.1	89.0	
55–64	29,409	16.0	91.0	
≥65	28,779	15.6	93.0	
Gender				
Male	74,323	40.4	89.7	0.04
Female	109,653	59.6	89.9	
Health insurance				
Commercial	134,732	73.2	89.1	ref.
Medicare	27,277	14.8	92.5	<0.001
Medicaid	21,540	11.7	91.4	<0.001
Unknown	433	0.25	84.4	0.4
Race/ethnicity				
White	103,169	56.1	95.0	ref.
American Indian	993	0.5	95.4	0.28
Asian	6,118	3.3	95.5	0.03
Black/African-American	12,943	7.0	95.5	<.01
Hispanic/Latino	3,278	1.8	94.1	0.01
Other/choose no answer	1,473	0.8	96.5	<0.01
No data	56,743	31.9	80.0	<0.001
Pregnant during 2006[b]	7,974	—	91.1	0.2
Matched nonpregnant	7,974		90.7	
Preferred language				
English	167,112	90.8	90.5	ref.
Other	9,898	5.4	92.4	<0.001
No data	6,972	3.7	73.6	<0.001
Tobacco use[c]				
Never used	103,248	56.1	—	—
Former user	46,361	25.2		
Current user	34,373	18.7		
Current user at all visits	31,771	17.3		
Successful quit	1,640	0.9		
Failed quit	839	0.5		
New user	123	1.0		

[a]Cochran–Armitage test for trend across age categories
[b]Cohort matched on age, race/ethnicity, and preferred language
[c]Tobacco use was calculated from all visit data. For pregnant women, smoking status was based on the 9-month pregnancy. For changes in status, at least two visits were required.

Table 4 demonstrates tobacco-use rates in selected populations with various combinations of characteristics in comparison to a common reference group. These predicted adjusted rates reflect the marginal impact of various combinations of demographic factors and illustrates some of the differences that currently exist between different population groups. They also illustrate the interactions between many of the demographic factors examined in this study; for example, men aged 18–24 years were particularly high users of tobacco. Rates of tobacco use among English-speaking men in this age group ranged from 14.7% (RR=0.64) among Asian men to 50% (RR=2.16) among American Indian men. Among non-English speakers, each of the groups were less likely to be tobacco users, although whites were most likely to be tobacco users (RR=0.53) while blacks were least likely (RR=0.32). Users of Medicaid, the proxy for lower SES, have elevated rates of tobacco use with relative rates ranging from 0.87 (Asians) to 2.92 (American Indians).

Discussion

This analysis demonstrates that EMR data can be used to measure tobacco-use prevalence among patients with a variety and combinations of characteristics. The sample sizes are large enough to identify prevalence rates for minority populations and demonstrate or adjust those rates for differential rates of various other characteristics within those populations. Important

Table 2. Tobacco use status by specific characteristic

Characteristics	Current user[a] (%)	Former user (%)	Never user (%)	*p* value
Age (years)[b]				
18–24	24.2	10.77	65.0	<0.001
25–34	22.8	15.5	61.7	
35–44	21.3	17.7	61.0	
45–54	20.9	25.5	53.5	
55–64	15.2	36.0	48.8	
≥65	7.7	42.9	49.4	
Overall (N=183,982)	18.7	25.2	56.1	
Gender				
Male	16.8	22.7	60.5	<0.001
Female	21.4	28.9	49.7	
Overall (*n*=183,976)	18.7	25.2	56.1	
Health insurance				
Commercial	17.3	23.7	59.0	ref.
Medicare	12.1	40.6	47.3	<0.001
Medicaid	34.3	14.1	51.6	<0.001
Unknown/self pay	16.0	33.5	50.5	<0.001
Overall (*n*=183,549)	18.7	25.2	56.1	
Race/ethnicity				
White	16.9	29.7	53.4	ref.
American Indian	40.0	29.4	30.6	<0.001
Asian	9.5	12.2	78.4	<0.001
Black/African-American	20.5	16.3	63.2	<0.001
Hispanic/Latino	17.8	23.4	58.8	0.09
Other	14.3	17.6	68.1	0.002
Overall (*n*=127,372)	17.1	27.2	55.7	
Pregnant in 2006	16.0	17.6	66.4	<0.001
Not pregnant[c]	16.9	23.1	60.0	
Overall (*n*=12,294)[c]	16.5	20.3	63.2	
Preferred language				
English	19.0	26.4	54.6	<0.001
Other	8.5	10.5	81.1	
Overall (*n*=177,010)	18.4	25.6	56.0	

[a]Defined as tobacco use anytime during the study time-frame
[b]Cochran–Armitage test for trend comparing tobacco users to non-users
[c]Cohort matched on age, race/ethnicity, and preferred language

to this conclusion is the analysis that tobacco status reporting rates differ very little from one category to another. This database has the potential to obtain relatively frequent repeat measures at low cost in order to track trends or to assess the effects of targeted interventions.

This capability requires that the medical group using the EMR must have made a similar effort to collect data that are usually not recorded or are recorded so inconsistently that they are unreliable, especially data regarding race/ethnicity, tobacco use, and language preference. The other variables should be reliably available in most systems, since they are required for normal care and billing. In this case, the effort to increase the rate of identification of race/ethnicity and language was part of an overall organizational goal of identifying any disparities by age, gender, or race/ethnicity for preventive services, vascular disease, diabetes, pregnancy, and asthma, and reducing them by 75% by 2010.

Although there are significant tobacco-use rates among most of the categories studied (except for pregnancy), this is primarily an artifact of the large sample sizes. None of the differences is large numerically except for the group with missing race/ethnicity data, and the majority of that group varies from the population average primarily by having a high never-use tobacco status.

The unadjusted prevalence rates reported in Table 2 differ from those found in the BRFSS system for Minnesota, which had such low sample sizes that it was necessary to combine 2005 and 2006 data in order to report on smoking status for racial/ethnic groups (apps.nccd.cdc.gov/brfss/). It "estimated" an overall cigarette use of 18.3%, with those aged 18–24 years smoking at 27.4% and race/ethnicity rates of 29.2% for African Americans, 19.8% for Hispanics, 18.2% for Asians, and 18.5% for whites (insufficient sample size for American Indian). In the current study, rates of tobacco use were 17.3% overall, 16.9% for whites, and 17.8% for Hispanics. The 1%–2% difference could in part reflect the more current data, but it may also be affected by the fact that BRFSS rates are adjusted to the entire state population and there was no such adjustment made in the current study. However, the rates for the other racial/ethnic groups were 9% lower for African Americans and Asians. In addition, the sample sizes are large enough that all but the Hispanic rates are statistically different from those of whites, whereas most of the BRFSS rates cannot be distinguished statistically.

The adjusted rates in Table 3 allow us to take advantage of the large sample sizes in order to demonstrate the impact of individual characteristics after controlling for the other characteristics in the table. They reveal a large difference between pregnant and nonpregnant women and generally heighten differences. Table 4's relative rates for combinations of characteristics also demonstrate the potential power of this system to identify more specific tobacco-use differences among people. For example, each of the race/ethnicity categories includes a heterogeneous collection of groups, especially recently arrived immigrant popu-

lations. For some of these groups, there may be sufficient sample size to explore tobacco-use rates by taking advantage of the information in the system about language preference and/or country of origin. It may also be possible to measure those rates by gender, since many immigrant groups are known to have large differences in tobacco use between men and women.

While the national surveys (NHIS, BRFSS, NAMCS) can provide some information about tobacco-use prevalence for some large subgroups of the population, they still do not include large enough sample sizes to study prevalence in populations with combinations of characteristics (e.g., for young male Hispanics). Most national surveys also have relatively low response rates, with potential differences between responders and nonresponders.[5,6] These disadvantages can be reduced in research studies that recruit and study specific groups in large numbers, but such studies of smoking prevalence are very expensive, are usually limited to a few groups, and are incapable of frequent repetition for surveillance and monitoring.[7–10] Studies of longitudinal cohorts like the Coronary Artery Risk Development in Young Adults (CARDIA) or the Framingham studies do provide such surveillance potential, but only for the particular subgroup or locality.[11–14]

Table 3. Comparison of subsample with complete data with whole population (n=126,553)

Characteristics	Current user[a](%)	Change in prevalence (%)	Former user (%)	Change in prevalence (%)
Age (years)[b]				
18–24	23.9	−0.3	11.6	0.8
25–34	21.3	−1.5	16.5	1.0
35–44	19.8	−1.5	19.0	1.3
45–54	19.2	−1.7	24.9	−0.6
55−64	13.9	−1.3	37.3	1.3
≥65	7.3	−0.4	43.9	1.0
Gender				
Male	19.0	−2.4	32.7	3.8
Female	15.8	−1.0	24.1	1.4
Health insurance				
Commercial	15.8	−1.5	25.6	1.9
Medicare	11.2	−0.9	41.9	1.3
Medicaid	33.3	−1.0	15.9	1.8
Unknown/self pay	15.1	−1.0	34.4	0.9
Race/ethnicity				
American Indian	40.0	0.0	29.3	−0.1
Asian	9.5	0.0	12.2	0.0
Black/African-American	20.5	0.0	16.4	0.1
Hispanic/Latino	17.8	0.0	23.4	0.0
White	16.9	0.0	29.8	0.1
Pregnant in 2006	15.9	0.1	18.4	0.8
Not pregnant[c]	15.8	1.1	24.6	1.5
Preferred language				
English	17.5	−1.5	28.2	1.8
Other	8.2	−0.3	11.3	0.8

[a]Defined as tobacco use anytime during the study time-frame
[b]Cochran–Armitage test for trend comparing tobacco users to non-users
[c]Cohort matched on age, race/ethnicity, and preferred language

Even though there are similarities between the population of this study and that of the state, they are not identical since nearly all those in the database of this study live in the metropolitan area and have some form of health insurance. In addition, the data are collected as part of medical care, so data are limited to those who visit for care and have more data on frequent visitors. It is also likely that the way questions are asked and answers are interpreted varies among different personnel. Patient reports may also differ from survey situations in the likelihood to falsify their tobacco use. Also, tobacco-use data are missing on 10% of the population and 80% of the visits, as well as race/ethnicity status on one third of the population. However, the proportion of data that is complete is increasing over time, and the organizational goal is to reach 100%. Correcting the missing tobacco-use data at visits should focus on those with a prior tobacco-user or recent-quitter status in order to maximize value from time-pressured staff. The decision to consider the 0.9% of users who later reported being quitters as current smokers might have introduced a slight overestimate of current prevalence, but counting them as quitters would have underestimated prevalence. Finally, these data will lack the power for estimating differences among small subgroups.

Nevertheless, this is a much larger data set than has been available for this type of surveillance of tobacco-use prevalence. If the proportion of visits at which tobacco-use status is recorded could be increased, it would even be possible to monitor rates and durations of quits and starts. This study demonstrates that such use of an electronic medical record can provide rapid and repeated prevalence assessments that allow us to track cohorts and to evaluate the effects of targeted interventions. It also allows analyses that combine the data in this report with healthcare utilization and cost data, at least for those healthcare services that are provided through this medical group. These capabilities should greatly strengthen

Table 4. Predicted rates of tobacco use for selected populations compared to a referent population

Description of population	N/*n* (126,553)	Current user (%)	Relative rate	Former user (%)	Relative rate
English-speaking white men aged 35–44 years with commercial insurance	1180	23.2	ref.	24.0	ref.
English-speaking men aged 18–24 years	105	50.0	2.16	15.3	0.66
American Indian					
Asian	782	14.8	0.64	7.6	0.33
Black	2069	29.2	1.26	10.2	0.44
Hispanic	421	27.6	1.19	15.0	0.65
White	8838	29.3	1.26	13.9	0.60
Medicaid insured					
American Indian	274	67.8	2.92	16.0	0.69
Asian	1409	20.1	0.87	13.9	0.60
Black	5287	32.1	1.38	13.4	0.58
Hispanic	671	34.0	1.46	22.4	0.97
White	5876	49.0	2.11	18.0	0.78
Pregnant women					
American Indian	75	29.8	1.28	27.5	1.18
Asian	614	4.6	0.20	11.8	0.51
Black	1398	9.9	0.43	13.3	0.57
Hispanic	317	10.9	0.47	23.7	1.02
White	3646	11.2	0.48	21.3	0.92
Non-English speakers					
Asian	2468	9.9	0.43	8.6	0.37
Black	2659	7.5	0.32	8.2	0.35
Hispanic	931	11.5	0.49	14.8	0.64
White	680	12.3	0.53	14.0	0.60
English-speaking women aged 18–24 years on Medicaid					
American Indian	81	71.5	3.08	10.1	0.44
Asian	492	19.7	0.85	8.4	0.36
Black	1171	33.2	1.43	8.2	0.35
Hispanic	271	36.1	1.56	14.3	0.62
White	6301	51.6	2.22	11.4	0.49
Non–English-speaking men aged 18–24 years on Medicaid					
Asian	18	12.4	0.53	3.6	0.16
Black	132	7.9	0.34	2.6	0.11
White	31	25.2	1.08	6.3	0.27

our ability to more effectively understand and reduce tobacco use.

Any public dissemination of information relating to the grant was made possible by Grant Number RC-2006-0053 from ClearWay Minnesota[SM]. The contents of this information are the sole responsibility of the authors and do not necessarily represent the views of ClearWay Minnesota[SM].

No financial disclosures were reported by the authors of this paper.

References

1. Chowdhury PP, Balluz L, Murphy W, et al. Surveillance of certain health behaviors among states and selected local areas—United States, 2005. MMWR Surveill Summ 2007;56:1–160.
2. Franks P, Fiscella K, Meldrum S. Racial disparities in the content of primary care office visits. J Gen Intern Med 2005;20:599–603.
3. NCHS. Design and estimation for the National Health Interview Survey, 1995–2004. Vital Health Stat 2000:1–31.
4. Hardy JR, Holford TR, Hall GC, Bracken MB. Strategies for identifying pregnancies in the automated medical records of the General Practice Research Database. Pharmacoepidemiol Drug Saf 2004;13:749–59.
5. Hill A, Roberts J, Ewings P, Gunnell D. Non-response bias in a lifestyle survey. J Public Health Med 1997;19:203–7.
6. Solberg LI, Hollis JA, Stevens VJ, Rigotti NA, Quinn VP, Aickin M. Does methodology affect the ability to monitor tobacco control activities? Implications for HEDIS and other performance measures. Prev Med 2003;37:33–40.
7. Levinson AH, Perez-Stable EJ, Espinoza P, Flores ET, Byers TE. Latinos report less use of pharmaceutical aids when trying to quit smoking. Am J Prev Med 2004;26:105–11.
8. Winkleby MA, Schooler C, Kraemer HC, Lin J, Fortmann SP. Hispanic versus white smoking patterns by sex and level of education. Am J Epidemiol 1995;142:410–8.
9. Maxwell AE, Bernaards CA, McCarthy WJ. Smoking prevalence and correlates among Chinese- and Filipino-American adults: findings from the 2001 California Health Interview Survey. Prev Med 2005;41:693–9.
10. Ward KD, Vander Weg MW, Kovach KW, et al. Ethnic and gender differences in smoking and smoking cessation in a population of young adult air force recruits. Am J Health Promot 2002;16:259–66.
11. Kiefe CI, Williams OD, Greenlund KJ, Ulene V, Gardin JM, Raczynski JM. Health care access and seven-year change in cigarette smoking. The CARDIA Study. Am J Prev Med 1998;15:146–54.
12. Kiefe CI, Williams OD, Lewis CE, Allison JJ, Sekar P, Wagenknecht LE. Ten-year changes in smoking among young adults: are racial differences explained by socioeconomic factors in the CARDIA study? Am J Public Health 2001;91:213–8.
13. Freund KM, D'Agostino RB, Belanger AJ, Kannel WB, Stokes JI. Predictors of smoking cessation: the Framingham Study. Am J Epidemiol 1992;135:957–64.
14. Wagenknecht LE, Craven TE, Preisser JS, Manolio TA, Winders S, Hulley SB. Ten-year trends in cigarette smoking among young adults, 1986–1996: the CARDIA study. Coronary artery risk development in young adults. Ann Epidemiol 1998;8:301–7.