Nicotine & Tobacco Research Advance Access published May 16, 2012

Original Investigation

Smoking-Attributable Mortality by Cause in the United States: Revising the CDC's Data and Estimates

Brian Rostron, Ph.D.

Center for Tobacco Products, Food and Drug Administration, Rockville, MD

Corresponding Author: Brian Rostron, Ph.D., *Center for Tobacco Products, Food and Drug Administration, 9200 Corporate Blvd, Room 300G Rockville, MD 20850, USA. Telephone: 301-796-9360; Fax: 240-276-3761; E-mail: brian.rostron@fda.hhs.gov*

Received December 31, 2011; accepted March 30, 2012

Abstract

Introduction: Smoking is the leading cause of preventable mortality in the United States, but the methods and data used in the Centers for Disease Control and Prevention's (CDC) published estimates of adult smoking-attributable mortality have not been substantially revised since their introduction in the 1980s.

Methods: We employed the CDC's general methodology for estimating smoking-attributable mortality but produced improved estimates by using recent, nationally representative relative risk data from the National Health Interview Survey—Linked Mortality Files and adjusting for confounding risk factors. We also produced estimates by smoking status and over time.

Results: Our use of more recent and nationally representative relative risks tended to decrease estimates of smoking deaths for men and increased estimates for women compared with the CDC's estimates. Adjustment for confounding factors further refned the estimates, particularly by smoking status. We estimated 200,000 smoking-attributable deaths for men and 180,000 smoking-attributable deaths for women in the United States in 2004. Estimated smoking-attributable mortality has fnally begun to decline for both U.S. men and women.

Conclusions: Our approach offers several substantive improvements in the estimation of smoking-attributable mortality by cause for the United States. Cigarette smoking remains a leading cause of preventable mortality in the United States, but we estimate that the number of smoking-attributable deaths has begun to decline.

Introduction

Cigarette smoking is the leading cause of preventable mortality in the United States (Centers for Disease Control and Prevention [CDC], 2002). The CDC periodically publishes estimates of smoking-attributable mortality for the U.S. adult population (CDC, 2005 , 2008), but the CDC has not substantially revised the methodology or data that are used in this procedure since their introduction in the 1980s (CDC, 1997; U.S. Department of Health and Human Services, 1989). The CDC's estimates are based on standard population-attributable risk calculations, with smoking prevalence data coming from the National Health Interview Survey (NHIS) and relative risk data coming from the American Cancer Society's Cancer Prevention Study II (CPS-II). The prevalence data come from the years for which smoking-attributable mortality is being estimated, and the relative risk data come from mortality follow-up in the CPS-II from 1982 to 1988. In many respects, these CPS-II relative risks from an earlier period are long overdue for replacement by more recent data. Doll and Peto (1981), for example, objected to the use of mortality risks for smokers from follow-up in the Cancer Prevention Study-I from 1959 to 1972 in the estimation of smoking-attributable mortality in the U.S. population in 1978, correctly noting that excess risk for smokers would have changed in the intervening period.

In this study, we address this problem by presenting updated and improved estimates of smoking-attributable mortality for U.S. adults. We employ the CDC's general methodology but use relative risks from more recent and nationally representative NHIS data in place of relative risks from CPS-II data. Researchers have previously shown the value of using NHIS data linked for mortality follow-up to estimate smoking-attributable mortality (Rogers, Hummer, Krueger, & Pampel, 2005; Rostron, 2011), and this study adds to the research feld with several important new contributions. First, we estimate smoking-attributable mortality by cause, whereas these previous studies have tended to use all-cause mortality. Second, we present comparable relative risks by smoking status and cause from CPS-II and NHIS data, showing the general changes in risks that have occurred over time. These changes may be due to factors such as changes in smoking duration and intensity among smokers as well as differences in characteristics such as socioeconomic status and health care among smokers compared with nonsmokers. Third, we present estimates of smoking-attributable morality for current and former smokers as well as estimates with and without adjustment for confounding risk factors such as body weight

doi: 10.1093/ntr/nts120
Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco 2012.

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

and alcohol consumption, which has not been done with the CDC methodology previously. Finally, we estimate smoking-attributable mortality for several years to show trends over time.

Methods

We used the CDC's method for estimating smoking-attributable mortality for U.S. adults as modifed for use with updated data and adjustment for confounding risk factors. The CDC's method estimates smoking-attributable mortality for persons aged 35 and over and is described in periodic reports (CDC, 2008) and on the CDC's Smoking-Attributable Mortality, Morbidity, and Economic Costs website (CDC, 2011a). The method estimates adult smoking-attributable mortality by sex as the sum of smoking-attributable mortality for 19 smoking-related causes. Smoking-attributable mortality for each cause is calculated by multiplying the number of deaths for the cause in the U.S. population by the smoking-attributable fraction (SAF) of deaths for the cause. The SAF is calculated as $SAF = ([p_1(RR_1 - 1) + p_2(RR_2 - 1)] / [p_1(RR_1 - 1) + p_2(RR_2 - 1) + 1])$, where p_1 is the proportion of current smokers, p_2 is the proportion of former smokers, RR_1 is the relative risk for current smokers relative to never-smokers, and RR_2 is the relative risk for former smokers relative to never-smokers. Smoking prevalence data come from the NHIS (National Center for Health Statistics [NCHS], 2011). Current smokers are defned as persons who report having smoked 100 or more cigarettes in their lives and who now smoke every day or some days. Former smokers have smoked at least 100 cigarettes but do not currently smoke. Relative risks come from the CPS-II and are estimated for ages 35 and over for most causes and for ages 35–64 and 65+ for ischemic heart disease and cerebrovascular disease. The SAF and smoking-attributable mortality for each cause are calculated for ages 35–64 and 65+ using the NHIS smoking prevalence data and national vital statistics deaths counts for these age groups as well as the CPS-II relative risks. The CDC's method does not adjust for confounding risk factors in the calculation of the CPS-II relative risks.

We updated the CDC's estimates by using relative risks from more recent and nationally representative NHIS data in place of CPS-II data. The CPS-II was a large longitudinal cohort study that was begun in 1982 (Thun et al., 1997). American Cancer Society volunteers recruited study participants, who were at least 30 years old, from households with at least one member who was 45 years old or over. CPS-II participants were more likely to be White, middle-class, and college-educated than the U.S. population of the time (Garfnkel, 1985; Thun et al., 1997). The NHIS is a nationally representative household health survey of the U.S. civilian noninstitutionalized population that is conducted on an annual basis by the CDC's NCHS (NCHS, 2011). NHIS interviewers personally survey approximately 35,000 households each year and one adult aged 18 and over from each household is selected to complete the Sample Adult Questionnaire, which includes detailed questions about smoking history. NHIS participants are followed for mortality through periodic linkage of their records with NCHS's National Death Index (MacMahon, 1983), which contains death certifcate information for all U.S. decedents since 1979, to produce the National Health Interview Survey—Linked Mortality Files (NHIS-LMF; NCHS, 2010). NCHS perturbs some date- and cause-of-death information in the public-use version of the NHIS-LMF in order to prevent identifcation of decedents but has shown that this perturbation has little effect on analyses conducted with the data (National Center for Health Statistics Data Linkage Team, 2010).

We used the public-use NHIS-LMF to calculate relative risks for selected causes by smoking status for NHIS participants from 1997 to 2004 with mortality follow-up through the end of 2006. A total of 242,937 adults aged 18 and over completed the sample adult questionnaire in these years and were eligible for mortality follow-up, and 241,016 of these individuals provided basic smoking status information. 17,030 deaths for this group were reported in the National Death Index, and 16,858 of these decedents had cause-of-death data available. We generally employed the CDC's 19 adult smoking-related causes of death, but combined some of the rarer causes in order to increase the accuracy and precision of our relative risk estimates, thus producing nine classes of causes (see for a side-by-side comparison of the two classifcation systems of smoking-related causes). We, along with the CDC, classified causes by their International Classifcation of Diseases, 10th Revision (ICD-10) codes. The NHIS-LMF does not include specifc cause-of-death data for two of the CDC causes—acute myeloid leukemia (ICD-10: C92.0) and chronic airway obstruction (ICD-10: J44)—so we used data for the comparable but more general causes of leukemia (ICD-10: C91–C95) and chronic lower respiratory diseases (ICD-10: J44, J47) in the estimation of relative risks. We conducted alternative analyses, such as calculating relative risks for other cancers omitting leukemia entirely, to examine the effects of these coding differences. The differences in overall estimates were minimal. We estimated relative risks for smoking status by cause as hazard ratios using Cox proportional hazard models with age as the time scale. Hazard ratios are similar to relative risks for the incidence levels and follow-up period in this analysis (Jewell, 2003). Following the CDC method, we estimated hazard ratios for ages 35 and over for most causes and for ages 35–64 and 65 and over for ischemic heart disease and cerebrovascular disease. We estimated hazard ratios both without ("age-adjusted") and with ("adjusted") an additional set of control variables that consisted of race/ethnicity, educational attainment, alcohol consumption, and body mass index (BMI). 4.7% of the survey participants did not have complete information for these control variables (most of the missing data was for BMI [3.0% of participants] and alcohol consumption [1.6% of participants]), and these participants were omitted from the adjusted analyses. Race/ethnicity was categorized as non-Hispanic White, non-Hispanic Black, non-Hispanic other race, and Hispanic. Educational attainment was categorized as less than high school graduate, high school graduate or equivalent, and more than high school graduate. Alcohol consumption was categorized as never drinker (consumed fewer than 12 drinks during lifetime), former drinker (consumed at least 12 drinks during lifetime but none during last year), light to moderate drinker (consumed 1–2 drinks on average on days consuming alcohol during last year for females, and 1–3 drinks for males), and heavy drinker (consumed 3+ drinks/day for females and 4+ for males). BMI was categorized as underweight or normal ($BMI < 25$), overweight ($25 \leq BMI < 30$), and obese ($BMI \geq 30$). Previous research (Rostron, 2011) has shown that relative risks calculated from these data are generally consistent during the follow-up period. We also evaluated the proportionality assumption of the hazard models through chi-squared tests of the scaled Rosenfeld residuals as a function of

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

Table 1. Deaths Among 1997–2004 National Health Interview Survey Adult Cohorts Through the End of 2006, by Sex, Smoking Status, and Cause

Smoking-related cause (ICD-10 codes)[a]	Males			Females		
	Never-smoker	Current smoker	Former smoker	Never-smoker	Current smoker	Former smoker
Upper aerodigestive cancer (C00–C15, C32)	31	55	62	13	11	13
Lung cancer (C33–C34)	55	326	328	89	264	208
Other cancer (C16, C25, C53, C64–C65, C67, C91–C95)	97	78	191	178	56	88
Ischemic heart disease (I20–I25)						
35–64 years	81	198	103	53	91	32
65+ years	438	240	792	954	182	402
Other heart disease (I00–I09, I26–I51)	152	99	238	465	80	181
Cerebrovascular disease (I60–I69)						
35–64 years	18	25	16	19	32	10
65+ years	113	44	165	394	70	136
Arterial disease (I70–I79)	30	28	56	70	35	38
Pneumonia, infuenza (J10–J18)	40	32	75	127	33	50
COPD (J40–J44, J47)	38	149	266	73	171	227
Total—smoking-related causes	1,093	1,274	2,292	2,435	1,025	1,385
Total—all deaths	2,089	2,053	3,743	4,749	1,692	2,427

Note. COPD = chronic obstructive pulmonary disease; ICD-10 = International Classifcation of Diseases, 10th Revision.
[a]Deaths are for ages 35 and over unless noted.

follow-up time. The proportionality assumption was generally upheld.

In our analysis using adjusted relative risks, we used the method presented by Flegal, Graubard, Williamson, and Gail (2005) for estimating deaths attributable to a risk factor while controlling for confounding risk factors. We estimated the mortality rate R for a cause by sex–age group as $R = I\sum r_i p_i$, where I is the baseline population mortality rate, r_i is the relative risk corresponding to each combination of risk factors in the adjusted analysis, and p_i is the prevalence of each combination of risk factors. We then estimated an alternative R^* as $R^* = I\sum r_i^* p_i$, where r_i^* is the alternative relative risk in which smoking level is set to the reference level but all other risk factors in the combination remain the same. We estimated the proportion of deaths attributable to smoking level as $(R - R^*)/R$. We then multiplied the resulting SAF of deaths by the number of deaths for the cause by sex–age group in the United States to estimate the number of smoking-attributable deaths by cause.

We produced estimates of smoking-attributable mortality using age-adjusted and adjusted NHIS-LMF relative risks for 2004, which is the most recent year for which the CDC has published estimates for the U.S. population (CDC, 2008, 2011a), for purposes of comparison. We also produced estimates using adjusted NHIS-LMF relative risks for 2000 and 2007 to show changes in smoking-attributable mortality over time.

We conducted all statistical and data analyses using R version 2.10.1 (R Development Core Team, 2009) and the survey package (Lumley, 2010). We used the appropriate NHIS-LMF sample weights and accounted for the NHIS complex survey design.

Results

Table 1 presents the number of deaths in the NHIS-LMF data for ages 35 and over by sex, smoking status, and cause category. Most causes have a reasonably large number of deaths for each smoker group. One exception is upper aerodigestive cancer deaths among women, given that this cause tends to be rare among women generally. Tables 2 and 3 present CPS-II relative risks, age-adjusted NHIS-LMF relative risks, and adjusted NHIS-LMF relative risks by smoking status for the smoking-related causes. Estimated NHIS-LMF relative risks for smokers tend to be lower than the corresponding CPS-II relative risks for upper aerodigestive and lung cancer among men and higher for lung cancer and chronic obstructive pulmonary disease among women. Adjustment for confounding risk factors tends to somewhat lower the estimated NHIS-LMF relative risks for male current smokers and raise the relative risks for female former smokers.

Tables 4 and 5 present the estimated number of smoking-attributable deaths by cause in the United States in 2004 using age-adjusted and adjusted NHIS-LMF relative risks, as well as the number of smoking-attributable deaths as estimated by the CDC's method. Use of age-adjusted NHIS-LMF relative risks in place of CPS-II relative risks for men increases estimates of smoking-attributable deaths among current smokers by 15% and decreases estimates of deaths among former smokers by

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

Table 2. Relative Risks for U.S. Males by Smoking Status and Cause From Selected Data Sources

CPS-II relative risks[a]			NHIS-LMF relative risks (age adjusted)					NHIS-LMF relative risks (adjusted)			
Smoking-related cause (ICD-10 codes)	Current smoker	Former smoker	Smoking-related cause (ICD-10 codes)	Current smoker	95% *CI*	Former smoker	95% *CI*	Current smoker	95% *CI*	Former smoker	95% *CI*
Malignant neoplasms			Malignant neoplasms								
Lip, oral, and pharynx (C00–C14)	10.89	3.40									
Esophagus (C15)	6.76	4.46	Upper aerodigestive (C00–C15, C32)	2.89	1.73–4.82	1.37	0.84–2.23	1.68	0.98, 2.90	1.19	0.71–1.97
Larynx (C32)	14.60	6.34									
Lung (C33–C34)	23.26	8.70	Lung (C33–C34)	13.83	9.90–19.44	3.97	2.88–5.48	11.71	8.30, 16.53	3.85	2.80–5.31
Stomach (C16)	1.96	1.47									
Pancreas (C25)	2.31	1.15									
Kidney (C64–C65)	2.72	1.73	Other (C16, C25, C64–C65, C67, C91–C95)	2.05	1.46–2.88	1.64	1.25–2.15	2.02	1.40, 2.92	1.58	1.18–2.13
Bladder (C67)	3.27	2.09									
Acute myeloid leukemia (C92.0)	1.86	1.33									
Circulatory diseases			Circulatory diseases								
Ischemic heart disease (I20–I25)			Ischemic heart disease (I20–I25)								
35–64 years	2.80	1.64	35–64 years	3.43	2.61–4.51	1.62	1.15–2.29	3.18	2.34, 4.33	1.59	1.11–2.27
65+ years	1.51	1.21	65+ years	2.16	1.81–2.57	1.19	1.04–1.36	1.96	1.62, 2.37	1.16	1.01–1.34
Other heart disease (I00–I09, I26–I51)	1.78	1.22	Other heart disease (I00–I09, I26–I51)	1.82	1.35–2.46	1.10	0.87–1.38	1.71	1.22, 2.39	1.09	0.85–1.41
Cerebrovascular disease (I60–I69)			Cerebrovascular disease (I60–I69)								
35–64 years	3.27	1.04	35–64 years	1.80	0.94–3.48	0.94	0.45–1.96	1.59	0.81, 3.11	1.07	0.50–2.26
65+ years	1.63	1.04	65+ years	1.22	0.80–1.86	0.96	0.73–1.26	1.08	0.71, 1.64	1.00	0.75–1.33
Athlerosclerosis (I70–I71)	2.44	1.33									
Aortic aneurysm (I71)	6.21	3.07	Arterial disease (I70–I79)	2.79	1.45–5.36	1.40	0.81–2.42	2.30	1.19, 4.45	1.40	0.84–2.36
Other arterial disease (I72–I79)	2.07	1.01									
Respiratory diseases			Respiratory diseases								
Pneumonia, infuenza (J10–J18)	1.75	1.36	Pneumonia, infuenza (J10–J18)	2.83	1.68–4.76	1.28	0.84–1.97	2.48	1.39, 4.43	1.29	0,80–2.05
Bronchitis, emphysema (J40–J42, J43)	17.10	15.64	Chronic obstructive pulmonary disease (J40–44, J47)	12.81	8.39–19.56	4.68	3.11–7.00	11.34	7.16, 17.96	4.99	3.24–7.71
Chronic airway obstruction (J44)	10.58	6.80									

Note. CPS-II = Cancer Prevention Study – II; ICD-10 = International Classifcation of Diseases, 10th Revision; NHIS-LMF = National Health Interview Survey—Linked Mortality Files.

[a]CPS-II data are from published CDC data (CDC, 2011a) and come from mortality follow-up from 1982 to 1988. NHIS-LMF data (National Center for Health Statistics, 2010) are for 1997–2004 NHIS adult cohorts followed for mortality through the end of 2006. Age-adjusted NHIS-LMF relative risks control for age. Adjusted NHIS-LMF relative risks control for age, race/ethnicity, educational attainment, alcohol consumption, and body mass index. Relative risks are for ages 35 and over unless noted.

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

Table 3. Relative Risks for U.S. Females by Smoking Status and Cause From Selected Data Sources

CPS-II relative risks[a]			NHIS-LMF relative risk (age adjusted)					NHIS-LMF relative risk (adjusted)			
Smoking-related cause (ICD-10 codes)	Current smoker	Former smoker	Smoking-related cause (ICD-10 codes)	Current smoker	95% *CI*	Former smoker	95% *CI*	Current smoker	95% *CI*	Former smoker	95% *CI*
Malignant neoplasms			Malignant neoplasms								
Lip, oral, and pharnyx (C00–C14)	5.08	2.29									
Esophagus (C15)	7.75	2.79	Upper aerodigestive (C00–C15, C32)	3.79	1.46–9.82	1.89	0.78–4.58	3.59	1.46–8.84	2.20	0.94–5.25
Larynx (C32)	13.02	5.16									
Lung (C33-C34)	12.69	4.53	Lung (C33–C34)	16.24	12.45–21.30	5.67	4.30–7.39	14.30	10.67–19.15	6.01	4.53–7.97
Stomach (C16)	1.36	1.32		1.60	1.16–2.21	1.47	1.09–1.98	1.88	1.34–2.64	1.72	1.24–2.37
Pancreas (C25)	2.25	1.55									
Kidney (C64–C65)	1.29	1.05	Other (C16, C25, C53, C64–C65, C67, C91–C95)								
Bladder (C67)	2.22	1.89									
Acute myeloid leukemia (C92.0)	1.13	1.38									
Cervix (C53)	1.59	1.14									
Circulatory diseases			Circulatory diseases								
Ischemic heart disease (I20–I25)			Ischemic heart disease (I20–I25)								
35–64 years	3.08	1.32	35–64 years	4.11	2.72–6.21	1.51	0.88–2.59	3.93	2.56–6.05	1.48	0.82–2.64
65+ years	1.60	1.20	65+ years	1.82	1.50–2.21	1.24	1.08–1.42	1.95	1.60–2.37	1.37	1.18–1.58
Other heart disease (I00–I09, I26–I51)	1.49	1.14	Other heart disease (I00–I09, I26–I51)	1.40	1.05–1.87	1.14	0.93–1.40	1.52	1.12–2.07	1.28	1.03–1.59
Cerebrovascular disease (I60–I69)			Cerebrovascular disease (I60–I69)								
35–64 years	4.00	1.30	35–64 years	3.72	2.01–6.89	1.84	0.78–4.32	3.39	1.81–6.33	2.07	0.85–5.07
65+ years	1.49	1.03	65+ years	2.12	1.60–2.80	0.99	0.78–1.26	2.11	1.59–2.81	1.09	0.85–1.41
Athlerosclerosis (I70–I71)	1.83	1.00									
Aortic aneurysm (I71)	7.07	2.07	Arterial disease (I70–I79)	3.71	2.23–6.16	1.65	1.00–2.73	2.98	1.71–5.19	1.52	0.92–2.52
Other arterial disease (I72–I79)	2.17	1.12									
Respiratory diseases			Respiratory diseases								
Pneumonia, infuenza (J10–J18)	2.17	1.10	Pneumonia, infuenza (J10–J18)	2.20	1.47–3.28	1.05	0.73–1.52	2.56	1.66–3.95	1.15	0.76–1.76
Bronchitis, emphysema (J40–J42, J43)	12.04	11.77	Chronic obstructive pulmonary disease (J40–J44, J47)	17.13	12.29–23.84	8.76	6.34–12.10	15.27	10.81–21.55	9.29	6.55–13.17
Chronic airway obstruction (J44)	13.08	6.78									

Note. CPS-II = Cancer Prevention Study—II; ICD-10 = International Classifcation of Diseases, 10th Revision; NHIS-LMF = National Health Interview Survey —Linked Mortality Files.

[a]CPS-II data are from published CDC data (CDC, 2011a) and come from mortality follow-up from 1982 to 1988. NHIS-LMF data National Center for Health Statistics, 2010) are for 1997–2004 NHIS adult cohorts followed for mortality through the end of 2006. Age-adjusted NHIS-LMF relative risks control for age. Adjusted NHIS-LMF relative risks control for age, race/ethnicity, educational attainment, alcohol consumption, and body mass index. Relative risks are for ages 35 and over unless noted.

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

26%. The overall effect is a decrease of 19,000 male deaths with the NHIS-LMF data. Use of age-adjusted NHIS-LMF relative risks in place of CPS-II relative risks for women increases estimates of smoking-attributable deaths among current smokers by 14%, and produces an overall increase of 17,000 deaths. Use of adjusted NHIS-LMF relative risks in place of age-adjusted NHIS-LMF relative risks for men produces a 4% decrease in smoking-attributable deaths among current smokers, and an overall decrease of 8,000 deaths among men. Use of adjusted NHIS-LMF relative risks in place of age-adjusted NHIS-LMF relative risks for women produces a 16% increase in smoking-attributable deaths among former smokers, and an overall increase of 10,000 deaths among women.

We also estimated smoking-attributable deaths over time using adjusted NHIS-LMF relative risks. Estimated smoking-attributable deaths among U.S. men consistently declined from 211,000 in 2000 to 200,000 in 2004 to 195,000 in 2007. Estimated smoking-attributable deaths among U.S. women also declined from 187,000 in 2000 to 180,000 in 2004 to 175,000 in 2007.

Discussion

We have produced improved estimates of smoking-attributable mortality for U.S. adults using the CDC's general methodology, and in doing so have provided several signifcant innovations. First, we have provided updated and nationally representative relative risks by smoking status for important smoking-related causes for the U.S. population. Previous research has suggested that CPS-II relative risks by smoking status for causes such as heart disease and stroke were generally similar to the comparable relative risks for these causes in the United States in the 1980s (Malarcher et al., 2000), but relative risks associated with smoking can change over time. For example, researchers observed substantial increases in relative risks among current smokers for all-cause and lung cancer mortality in the CPS-II, with mortality follow-up from 1982 to 1988, compared with the CPS-I, with mortality follow-up from 1959 to 1965 (Thun et al., 1997). These increases were most dramatic among women. Results presented here suggest that relative risks for causes such as lung cancer and Chronic Obstructive Pulmonary Disease may have continued to rise among female smokers while declining among male smokers. Changes over time in smoking duration and intensity may have contributed to these trends, given that the change in risk is most apparent with lung cancer. Estimated relative risks for male smokers were approximately double those of female smokers in the CPS-II, but the estimated risks for female smokers in the NHIS-LMF were actually higher than those for males. The convergence of these estimated risks is consistent with evidence that female smokers came to more closely resemble male smokers over time in terms of their smoking behavior (Forey, Hamling, Lee, & Wald, 2002). We also found much lower estimated relative risks for upper aerodigestive cancer in the NHIS-LMF data compared with CPS-II data, particularly among men. Estimates from the NHIS-LMF for this cause are

Table 4. Smoking-Attributable Deaths for U.S. Males by Smoking Status and Cause, 2004

	Estimates using age-adjusted NHIS-LMF relative risks[a]			Estimates using adjusted NHIS-LMF relative risks		
Smoking-related cause (ICD-10 code)	Current smoker deaths	Former smoker deaths	Total smoker deaths	Current smoker deaths	Former smoker deaths	Total smoker deaths
Upper aerodigestive cancer (C00–C15, C32)	3,503	2,069	5,572	2,204	1,174	3,378
Lung cancer (C33-C34)	35,981	31,373	67,354	36,328	30,669	66,997
Other cancer (C16, C25, C64–C65, C67, C92.0)	4,254	8,902	13,156	4,368	8,381	12,749
Ischemic heart disease (I20–I25)						
35–64 years	18,640	5,814	24,454	18,672	5,229	23,901
65+ years	15,228	14,911	30,139	13,657	13,294	26,951
Other heart disease (I00–I09, I26–I51)	6,260	3,049	9,309	5,924	2,912	8,836
Cerebrovascular disease (I60–I69)						
35–64 years	1,628	0	1,628	1,362	135	1,497
65+ years	958	0	958	332	76	408
Arterial disease (I70–I79)	2,403	2,373	4,776	2,007	2,523	4,530
Pneumonia, infuenza (J10–J18)	3,851	2,807	6,658	3,412	2,825	6,237
COPD (J40–J44)	17,357	25,817	43,174	17,373	26,722	44,095
Total	110,063	97,115	207,178	105,639	93,940	199,579
CDC total[b]	95,970	130,498	226,468			

Note. CDC = Centers for Disease Control and Prevention; COPD = chronic obstructive pulmonary disease; CPS-II = Cancer Prevention Study II; ICD-10 = International Classification of Diseases, 10th Revision; NHIS-LMF = National Health Interview Survey—Linked Mortality Files.

[a]Estimates with NHIS-LMF relative risks were calculated using the methods described in the text.

[b]CDC totals were calculated using the CDC's method (CDC, 2011a) for estimating smoking-attributable mortality, which uses age-adjusted CPS-II relative risks.

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

Table 5. Smoking-Attributable Deaths for U.S. Females by Smoking Status and Cause, 2004

Smoking-related cause (ICD-10 code)	Estimates using age-adjusted NHIS-LMF relative risks[a]			Estimates using adjusted NHIS-LMF relative risks		
	Current smoker deaths	Former smoker deaths	Total smoker deaths	Current smoker deaths	Former smoker deaths	Total smoker deaths
Upper aerodigestive cancer (C00–C15, C32)	1,296	968	2,264	1,306	1,131	2,437
Lung cancer (C33–C34)	29,493	22,103	51,596	28,020	23,040	51,060
Other cancer (C16, C25, C53, C64–C65, C67, C92.0)	2,188	3,954	6,142	2,778	5,156	7,934
Ischemic heart disease (I20–I25)						
35–64 years	7,607	1,334	8,940	7,775	1,102	8,877
65+ years	12,667	12,325	24,992	12,610	16,491	29,101
Other heart disease (I00–I09, I26–I51)	3,370	3,434	6,804	3,773	6,200	9,973
Cerebrovascular disease (I60–I69)						
35–64 years	3,331	1,100	4,431	3,327	1,184	4,511
65+ years	7,675	0	7,675	7,187	1,978	9,165
Arterial disease (I70–I79)	3,394	2,447	5,841	2,715	2,050	4,765
Pneumonia, infuenza (J10–J18)	3,366	420	3,786	3,745	1,102	4,847
COPD (J40–J44)	20,012	27,859	47,871	19,038	28,551	47,589
Total	94,399	75,944	170,343	92,274	87,985	180,259
CDC total[b]	82,935	70,320	153,255			

Note. CDC = Centers for Disease Control and Prevention; COPD = chronic obstructive pulmonary disease; CPS-II = Cancer Prevention Study II; ICD-10 = International Classifcation of Diseases, 10th Revision; NHIS-LMF = National Health Interview Survey —Linked Mortality Files.

[a]Estimates with NHIS-LMF relative risks were calculated using the methods described in the text.

[b]CDC totals were calculated using the CDC's method (CDC, 2011a) for estimating smoking-attributable mortality, which uses age-adjusted CPS-II relative risks.

similar to those produced by a meta-analysis of studies of upper digestive tract cancer (ICD-10: C10–C15), which found pooled relative risks of 3.57 for current smokers (95% *CI* = 2.63–4.84) and 1.18 for former smokers (95% *CI* = 0.73–1.91) compared with never-smokers. In addition to changes in smoking intensity and duration, other factors that could affect relative risks include changes in the relative socioeconomic status and health care of smokers compared with nonsmokers and secular trends in disease risk by cause among nonsmokers. For example, smoking has become increasingly concentrated among individuals with lower educational attainment in the United States over time (Dube, Asman, Malarcher, & Caraballo, 2009).

Second, we have used these recent relative risks to produce improved estimates of smoking-attributable mortality by cause that are presented by smoking status and adjust for confounding factors. The CDC does not present estimates of smoking-attributable mortality by smoking status and does not adjust for confounding factors in its calculation of relative risks. Previous research (Thun, Apicella, & Henley, 2000) has shown that such adjustment does not have much effect on overall estimates of smoking-attributable mortality. Our results generally support this conclusion. For example, we estimated 380,000 overall smoking-attributable deaths in 2004 using adjusted NHIS-LMF relative risks compared with 378,000 deaths using age-adjusted NHIS-LMF relative risks. On the other hand, research has indicated that adjustment for confounding factors does generally have some effect on estimates by sex and smoking status (Rostron, 2011; Thun et al., 2000). In particular, adjustment tends to decrease estimates for male current smokers and increase estimates for female former smokers. We also found these same general results with our estimates. The difference in estimates with and without adjustment for confounding factors was most pronounced for female former smokers at 16%.

The differences produced by our use of more recent relative risks and adjustment for confounding factors are particularly evident when comparing estimates of smoking-attributable mortality by sex. The CDC estimates that the sex ratio of adult smoking-attributable deaths for U.S. males compared with females in 2004 was 1.48. Our estimate of the sex ratio using unadjusted NHIS-LMF relative risks is 1.22. Much of this decrease is due to the substantial reduction in deaths among male former smokers as estimated by NHIS-LMF data compared with CPS-II data, presumably due to earlier cessation among male smokers over time. The estimated sex ratio with adjusted NHIS-LMF relative risks is 1.11. In general, our estimates indicate that the magnitude of smoking-attributable mortality for U.S. men and women is much closer than the CDC's estimates would suggest. This result is consistent with estimates from other methods (Oza, Thun, Henley, Lopez, & Ezzati, 2011; Peto, Lopez, Boreham, & Thun, 2006) that fnd that smoking-attributable mortality in the United States is about 10%–20% higher for men than women. Once again, these results are consistent with evidence that smoking prevalence, duration, and intensity among U.S. women came to more closely resemble the smoking behavior of U.S. men over time (Forey et al., 2002).

Our estimates of smoking-attributable mortality estimated by cause are lower than some estimates previously obtained from more inclusive cause categories (Peto et al., 2006) or all-cause mortality (Rostron, 2011). This difference in estimates from all-cause and cause-specifc mortality has been observed previ - ously with United States (Burns, Garfnkel, & Samet, 1997) and

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

international data (Katanoda et al., 2008). To some extent, this difference may occur because relative risks for all-cause mortality for smokers may include some excess mortality risk that is not due to smoking, even with extensive controls for confounding risk factors. The difference may also occur because of the existence of additional causes of death with appreciable smoking-attributable mortality that is not included in cause-specifc estimates. For example, hypertensive disease (ICD-10: I10–I13) is not included in the CDC's method even though the relative risks for this cause in the NHIS-LMF data are 2.13 (95% *CI* = 1.23–2.70) for female current smokers and 1.41 (95% *CI* = 0.96–2.07) for female former smokers.

Third, we have estimated smoking-attributable mortality over time and found that the number of estimated deaths has peaked and fnally begun to decline for both men and women in the United States. This result is consistent with estimates for men from some other estimation methods, although it is generally a new observation for women. Estimates from the indirect method of Peto et al. (2006), for example, suggest that smoking-attributable mortality peaked among U.S. men between 1990 and 1995, although these estimates suggest that smoking-attributable mortality among U.S. women was still increasing as of 2000. Some of the decline in estimated smoking-attributable mortality was due to decreases in smoking prevalence in the United States over time. The prevalence of current smokers among persons aged 35–64, for example, declined from 2000 to 2007 from 28% to 23% for men and from 24% to 20% for women (NCHS, 2011). The remainder was due to signifcant decreases in deaths from important smoking-related causes. Deaths from ischemic heart disease for persons aged 65 and over decreased from 201,000 in 2000 to 158,000 in 2007 for U.S. men, and from 233,000 in 2000 to 170,000 in 2007 for U.S. women (CDC, 2011b). Our method could be further enhanced through the development of methods for constructing *CI*s for the estimates, analogous to methods for similar estimates of population-attributable risk (Flegal et al., 2005; Graubard & Fears, 2005). These methods would take into account variability in estimates at a particular point in time due to imprecision in estimates of smoking prevalence and relative risks as well as variability in estimates over time due to imprecision in estimates of smoking prevalence.

Finally, we have shown multiple advantages of using NHIS-LMF data in place of CPS-II data in the estimation of smoking-attributable mortality for the United States and other populations. The CPS-II data have the advantage of a larger sample size than the NHIS-LMF data, which may allow for the estimation of relative risks for more specific and rarer causes. The NHIS-LMF, on the other hand, has the advantage of being regularly updated with recent cohorts from the NHIS and periodic linkages for mortality follow-up with the National Death Index. All of these data sources are produced and maintained by the CDC's NCHS, and the CDC should consider use of this data source in its estimation of U.S. smoking-attributable relative risks and mortality. CPS-II relative risks are also used in the indirect method (Peto, Lopez, Boreham, Thun, & Heath, 1992) to estimate smoking-attributable mortality for other national populations (Peto et al., 2006). The results presented here suggest that consideration should be given to using updated NHIS-LMF relative risks with this method as well.

The estimates presented here illustrate the continuing magnitude of mortality caused by cigarette smoking in the United States. They also show the effects of changes in smoking behavior over time, as estimates of smoking-attributable mortality among U.S. women now almost equal estimates for men. Fortunately, the estimates also suggest that smoking-attributable mortality has begun to decline for men and women in the United States, indicating that the mortality burden from cigarettes should decline over time.

Funding

The author is employed by the Food and Drug Administration.

Declaration of Interests

The author has no competing interests.

Acknowledgments

This paper refects the views of the author and should not be construed to refect FDA's views or policies.

References

Burns, D. M., Garfnkel, L., & Samet, J. M. (1997). Introduction, Summary, and Conclusions. *In Smoking and tobacco control monograph 8: Changes in cigarette-related disease risks and their implications for prevention and control* (pp. 1–11). Bethesda, MD: National Cancer Institute.

Centers for Disease Control and Prevention. (1997). Smoking-attributable mortality and years of potential life lost—United States, 1984. *Morbidity & Mortality Weekly Report*, *46*, 444–451.

Centers for Disease Control and Prevention. (2002). Smoking-attributable mortality, years of potential life lost, and economic costs—United States, 1995–1999. *Morbidity & Mortality Weekly Report*, *51*, 300–303.

Centers for Disease Control and Prevention. (2005). Annual smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 1997–2001. *Morbidity & Mortality Weekly Report*, *54*, 625–628.

Centers for Disease Control and Prevention. (2008). Smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 2000–2004. *Morbidity & Mortality Weekly Report*, *57*, 1226–1228.

Centers for Disease Control and Prevention. (2011a). *Smoking-Attributable Mortality, Morbidity, and Economic Costs (SAMMEC)*. Retrieved from http://apps.nccd.cdc.gov/sammec

Centers for Disease Control and Prevention. (2011b). *WONDER*. Retrieved from http://wonder.cdc.gov/

Doll, R., & Peto, R. (1981). The causes of cancer: Quantitative estimates of avoidable risks of cancer in the United States today. *Journal of the National Cancer Institute*, *66*, 1191–1308.

Dube, S. R., Asman, K., Malarcher, A., & Caraballo, R. (2009). Cigarette smoking among adults and trends in smoking cessation—

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012

United States, 2008. *Morbidity & Mortality Weekly Report*, *58*, 1227–1232.

Flegal, K., Graubard, B., Williamson, D., & Gail, M. (2005). Excess deaths associated with underweight, overweight, and obesity. *Journal of the American Medical Association*, *293*, 1861–1867. doi:10.1001/jama.293.15.186

Forey, B., Hamling, J., Lee, P., & Wald, N. (2002). *International Smoking Statistics* (2nd ed.) Oxford: Oxford University Press.

Garfnkel, L. (1985). Selection, follow-up, and analysis in the American Cancer Society prospective studies. *National Cancer Institute monograph*, *67*, 49–52.

Graubard, B. I., & Fears, T. R. (2005). Standard errors for attributable risk for simple and complex sample designs. *Biometrics*, *61*, 847–855. BIOM355; doi:10.1111/j.1541-0420.2005.00355.x

Jewell, N. P. (2003). *Statistics for epidemiology*. Boca Raton, FL: Chapman and Hall/CRC.

Katanoda, K., Marugame, T., Saika, K., Satoh, H., Tajima, K., Suzuki, T., et al. (2008). Population attributable fraction of mortality associated with tobacco smoking in Japan: A pooled analysis of three large-scale cohort studies. *Journal of Epidemiology*, *18*, 251–264. doi:10.2188/jea.JE2007429

Lumley, T. (2010). *Survey: Analysis of complex survey samples (Version 3.23)*. Seattle, WA: Thomas, Lumley.

MacMahon, B. (1983). The national death index. *American Journal of Public Health*, *73*, 1247–1248. doi:10.2105/AJPH.73.11.1247

Malarcher, A. M., Schulman, J., Epstein, L. A., Thun, M. J., Mowery, P., Pierce, B., et al. (2000). Methodological issues in estimating smoking-attributable mortality in the United States. *American Journal of Epidemiology*, *152*, 573–584. doi:10.1093/aje/152.6.573

National Center for Health Statistics. (2010). *NHIS linked mortality fles* . Retrieved from http://www.cdc.gov/nchs/data_access/data_linkage/mortality/nhis_linkage.htm

National Center for Health Statistics. (2011). *National Health Interview Survey*. Retrieved from http://www.cdc.gov/nchs/nhis.htm

National Center for Health Statistics Data Linkage Team. (2010). *Comparative analysis of the NHIS public-use and restricted-use linked mortality fles: 2010 Data Release*. Retrieved from http://www.cdc.gov/NCHS/data/datalinkage/nhis_mort_compare_2010_fnal.pdf

Oza, S., Thun, M. J., Henley, S. J., Lopez, A. D., & Ezzati, M. (2011). How many deaths are attributable to smoking in the United States? Comparison of methods for estimating smoking-attributable mortality when smoking prevalence changes. *Preventive Medicine*, *52*, 428–433. doi:S0091-7435(11)00160-5;10.1016/j.ypmed.2011.04.007

Peto, R., Lopez, A., Boreham, J., & Thun, M. (2006). *Mortality from smoking in developed countries 1950–2000*. Retrieved from http://www.ctsu.ox.ac.uk/deathsfromsmoking/

Peto, R., Lopez, A. D., Boreham, J., Thun, M., & Heath, C., Jr. (1992). Mortality from tobacco in developed countries: Indirect estimation from national vital statistics. *The Lancet*, *339*, 1268–1278. doi:0140–6736.(92)91600-D

R Development Core Team. (2009). *R: A language and environment for statistical computing (Version 2.10.1)*. Vienna, Austria: Author.

Rogers, R., Hummer, R., Krueger, P., & Pampel, F. (2005). Mortality attributable to smoking in the United States. *Population and Development Review*, *31*, 259–292. doi:10.1111/j.1728-4457.2005.00065.x

Rostron, B. (2011). Smoking-attributable mortality in the United States. *Epidemiology*, *22*, 350–355. doi:10.1097/EDE.0b013e3182126729

Thun, M. J., Apicella, L. F., & Henley, S. J. (2000). Smoking vs other risk factors as the cause of smoking-attributable deaths: Confounding in the courtroom. *Journal of the American Medical Association*, *284*, 706–712. doi:joc00088

Thun, M., Day-Lally, C., Myers, D., Calle, E., Flanders, W., Zhu, B.-P., et al. (1997). Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). *In Smoking and tobacco control monograph 8: changes in cigarette-related disease risks and their implications for prevention and control*. (pp. 305–382). Bethesda, MD: National Cancer Institute.

U.S.Department of Health and Human Services. (1989). *Reducing the Health Consequences of Smoking: A Report of the Surgeon General*. Rockville, MD: Centers for Disease Control, Offce on Smoking and Health.

Downloaded from http://ntr.oxfordjournals.org/ at Society for Research on Nicotine and Tobacco member access on May 17, 2012