Cancer Mortality Among a Representative Sample of Nonsmokers in the United States During 1966-68 [1]

James E. Enstrom, *Ph.D.*, [2] *and* Frank H. Godley, *Ph.D.* [3]

ABSTRACT—Data are presented on cancer and total mortality among a representative sample of nonsmokers and the total population 35-84 years of age in the United States during 1966-68 that measured the influence of cigarette smoking on mortality rates, independent of other health-related factors. Of all U.S. white males, those who never smoked cigarettes have a total age-adjusted cancer death rate which is 37% less than that of males as a whole and 53% less than that of those who currently smoke cigarettes. Correspondingly, of all U.S. white females, those who never smoked cigarettes have a total age-adjusted cancer death rate which is 15% less than females as a whole and 33% less than that of those who currently smoke cigarettes. The largest cancer rate reduction in the nonsmokers is concentrated in the respiratory system. Nonsmokers have an age-adjusted total death rate which is about 20% less than the population as a whole and about 43% less than current cigarette smokers. These and other results and methodologic issues are discussed.—JNCI 65: 1175-1183, 1980.

Previous studies have established a strong, positive correlation between cigarette smoking and mortality, with cancer mortality showing some of the highest correlations. Beginning in 1964, the *Smoking and Health Report* of the Surgeon General has summarized the evidence (*1*). Only one previous study in the United States has been based on representative national samples and that was limited to lung cancer deaths during 1958-59 (*2, 3*).

The findings here are based on a 1966-68 survey that measures among a representative U.S. sample the effect of smoking status on all causes of death independent of other health-related factors. The nonsmokers in this survey are meant to be used as a reference for comparisons with other nonsmoking populations who may be subject to additional health-related influences and selection biases. Our focus is on cancer mortality among nonsmokers and their risks relative to those experienced by the total population. "Nonsmokers" here are defined as persons who smoked fewer than 5 packs of cigarettes (100 cigarettes) in their lifetimes. However, they include former and current cigar and/or pipe smokers.

Unlike most other low-risk populations, nonsmokers comprise a substantial proportion of the adult population. They are not limited to a particular geographic region, ethnic group, or religion. Moreover, our use of national probability samples to examine the mortality experience of nonsmokers minimizes selection biases with respect to other environmental, socioeconomic, and life-style factors.

MATERIALS AND METHODS

Investigators who conducted most previous studies of cigarette smoking and mortality used a single-sample, prospective design. Mortality ratios were obtained from the relative proportions of the original sample who died within a specified follow-up period and then were weighted by survival times. However, serious disadvantages associated with representativeness and problems of follow-up are evident in the prospective studies (*1*). Nonresponse may have been as much as 32% overall and even greater among smokers. The bias in the mortality ratios of the respondents was estimated to be about a 0.2 or 0.3 overstatement because of lower mortality ratios of the nonrespondents. The low level of nonresponse in the present study (5%) indicated that nonresponse was negligible as a source of bias.

Our approach was to use a cross-sectional, two-sample design to estimate the relative risk of nonsmokers' mortality. This study was analogous to the one of representative samples of 1958-59 lung cancer deaths among U.S. whites (*2, 3*). One sample was representative of U.S. deaths during 1966-68, and the other resembled the general U.S. population during the same period. The smoking characteristics of both samples were determined and then mortality rates were calculated as a function of smoking status. Some limited results of the study, described in detail elsewhere,[4] have been published (*4-6*).

Sources of data.—Data were from two sources. Estimates of observed deaths of nonsmokers 35-84 years of age in the United States in 1966-68 were based on the NMS, a follow-back survey linked to a probability sample of 19,526 death registration records that in-

ABBREVIATIONS USED: CPS = Current Population Survey(s); ICD = International Classification of Diseases; NCHS = National Center for Health Statistics; NMS = National Mortality Survey; SMR = standardized mortality ratio(s).

[1] Presented at the Workshop on Populations at Low Risk of Cancer conducted at Snowbird, Utah, August 23-25, 1978.

[2] School of Public Health and Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Calif. 90024.

[3] Division of Analysis, National Center for Health Statistics, 5700 East-West Highway, Hyattsville, Md. 20782.

[4] Godley, FH. Cigarette smoking, social factors, and mortality: New estimates from representative national samples. Unpublished Ph.D. dissertation, University of Maryland, College Park, Md., 1974.

cluded deaths of 11,318 white males and 5,636 females.[5] The underlying cause of each death was assigned in accordance with the Seventh Revision of the ICD (*7, 8*). Questionnaires were mailed to surviving family members and others named on the death certificate who provided smoking histories and social characteristics of the deceased. The second source of data is the CPS conducted by the U.S. Bureau of the Census in August 1967 (*9*). Smoking and socioeconomic information comparable to that for the decedent was obtained by means of household interviews for a probability sample of 60,920 adults 35-84 years of age, including 25,266 white males and 29,308 white females.[4] The technical details of each survey are given in the Appendix. Also included there are the age-sex distributions of the deaths and population-at-risk by smoking status and a sample calculation of the weighted number of deaths when the sample number of deaths is used.

Inferences that can be drawn from the data in this study are conditioned by certain features of its design, sampling, and data-collection methods, as discussed in detail in the Appendix. In prospective studies, the decedents are a subset of the total study population; examples are the American Cancer Society and U.S. Veterans cohort studies in which a beginning sample is initially questioned regarding both current and past smoking habits (*10, 11*). The decedents of a prospective study, who are in the original sample, match exactly with an individual in the study population. The cross-sectional design assumes statistical matching between the two independently drawn samples: decedents and population. This problem is common to conventional vital statistics rates and must be balanced with the expense, time, and follow-up problems of prospective studies. The value of a two-sample cross-sectional design was demonstrated in national studies of smoking and lung cancer mortality during 1958-59 (*2, 3*).

Definitions.—Detailed estimates of cigarette smoking history by amount smoked and other relevant smoking variables were not made. Instead, persons were simply classified by lifetime history of having "never smoked" or "ever smoked" cigarettes (at least 5 packs or 100 cigarettes). "Current" cigarette smokers are persons who have smoked at least 100 cigarettes in their lifetimes and have smoked cigarettes during the past 12 months. We coordinated smoking questions in the supplement to the CPS with those in the death follow-up survey in an effort to gain comparability. Unlike sizable proportions of the CPS sample who provided data on their smoking habits and other personal characteristics, such data were not available for the NMS death sample. Moreover, a mailed questionnaire was used in the NMS, whereas a household interview was used by the CPS for collection of data on the living population. Evidence from another survey indicated that agreement between the smoking history reported for a decedent by next of kin and by his own reports when alive was about 90%. Moreover, Haenszel et al. (*2*) noted that this type of bias is largest with respect to medium and heavy smokers but virtually nonexistent for nonsmokers.

RESULTS

Table 1 contains the average annual age-specific total death rates for 1966-68 for U.S. white males and females by smoking status: never smoked cigarettes, currently smoke cigarettes, and the total sample. The rates are given in 10-year intervals for adults 35-84 years old, with an overall rate age adjusted to the 1940 U.S. population. The 1940 U.S. population is the standard reference used by the NCHS (*7*). Included in table 1 are age-specific total death rates for all U.S. whites based on the NMS and CPS and the average total death rates as published in the 1966-68 and 1970 Vital Statistics of the United States (*8*). The NMS-CPS rates generally agree with the 1966-68 U.S. rates to within 3%, except in the 75- to 84-year-old category, in which they are up to 10% higher than the U.S. rates. (The CPS includes only the noninstitutionalized population as explained in the Appendix.) Agreement between the survey rates and the U.S. vital statistics is good. Finally, it is evident that the nonsmokers in this survey have significantly lower rates than those of the total population, which conforms with previous epidemiologic studies of nonsmokers (*1, 10, 11*). Nonsmokers have an age-adjusted total death rate about 20% less than that of the population as a whole and about 43% less than the rate for current cigarette smokers.

In addition to the 1966-68 average annual age-specific cancer death rates for U.S. white males and females by smoking status, table 2 includes the rates for total cancer and for cancers of the lung, breast, and prostate, all of which are based on more than 100 sample deaths. The U.S. white males who never smoked cigarettes have a total age-adjusted cancer death rate 37% less than that of U.S. white males as a whole and 53% less than that of white males who are currently cigarette smokers. With the rates for total cancer based on the NMS and CPS are those given in the 1966-68 and 1970 U.S. vital statistics. The agreement between the survey and the vital statistics rates is within 11% for those less than 65 years old. Over 65 years, the survey rates are 7-17% higher than those for vital statistics. These differences are greater than the corresponding differences for total mortality primarily because of statistical fluctuation in the sampling of cancer deaths in the NMS.

The 1966-68 weighted numbers of observed cancer deaths for U.S. whites who never smoked cigarettes for the sites listed based on at least 10 sample deaths in males and/or females are given in table 3. Also included are the expected numbers of cancer deaths

[5] National Center for Health Statistics. 1966-68 National Mortality Survey. Unpublished technical notes and computer tapes, 1976.

TABLE 1.—*The 1966–68 average annual age-specific total death rates (deaths/100,000) among U.S. whites by age, sex, and cigarette smoking status*

Age, yr	U.S. whites who never smoked cigarettes[a]	U.S. whites who currently smoke cigarettes[a]	All U.S. whites		
			1966–68 NMS + CPS[a]	1966–68 U.S. vital statistics	1970 U.S. vital statistics
Males					
35–44	216	442	352	342	344
45–54	570	1,183	937	904	883
55–64	1,686	2,878	2,352	2,246	2,203
65–74	3,654	6,572	4,933	4,968	4,810
75–84	9,528	13,930	10,749	9,918	10,099
Total: 35–84[b]	1,425	2,459	1,918	1,852	1,827
Females					
35–44	152	245	201	195	193
45–54	333	597	464	462	463
55–64	771	1,444	1,015	1,022	1,015
65–74	2,398	4,129	2,702	2,628	2,471
75–84	7,044	13,123	7,600	6,890	6,699
Total: 35–84[b]	894	1,606	1,058	1,017	989

[a] Death rates are based on the 1966–68 NMS and August 1967 CPS.
[b] Rates are age adjusted by the direct method to the 1940 U.S. population 35–84 yr old.

based on the death rates among all the whites in the 1966-68 NMS sample. With this group as the reference population, the total cancer SMR for whites were 67 and 87% for male and female nonsmokers, respectively. Using 1970 U.S. whites as the reference population, as was done in a previous publication summarizing some of these NMS results (5), we determined that the total cancer SMR were 70% for white

TABLE 2.—*The 1966–68 average annual age-specific cancer death rates (deaths/100,000) among U.S. whites by age, sex, and cigarette smoking status*[a]

Age, yr	Lung cancer			U.S. whites who never smoked cigarettes[b]	Total cancer				
	U.S. whites who never smoked cigarettes[b]	U.S. whites who currently smoke cigarettes[b]	All U.S. whites[b]		U.S. whites who never smoked cigarettes[b]	U.S. whites who currently smoke cigarettes[b]	All U.S. whites		
							1966–68 NMS + CPS[b]	1966–68 U.S. vital statistics	1970 U.S. vital statistics
Males									
35–44	2.3 (2)	17.5 (38)	12.9 (48)	0.7 (1)	32.0 (28)	56.8 (125)	49.2 (185)	50	50
45–54	3.5 (3)	81.1 (163)	59.3 (213)	2.1 (3)	80.7 (42)	211.2 (340)	172.4 (476)	170	172
55–64	33.4 (24)	270.4 (365)	177.6 (493)	15.2 (18)	306.0 (102)	685.7 (646)	527.5 (946)	479	498
65–74	63.8 (43)	533.9 (298)	288.5 (499)	79.9 (89)	655.5 (204)	1,531.3 (543)	1,079.7 (1,041)	985	997
75–84	87.9 (36)	746.1 (122)	267.2 (211)	244.4 (165)	1,360.8 (294)	2,224.3 (262)	1,548.0 (760)	1,472	1,593
Total: 35–84[c]	19.9 (108)	179.8 (986)	103.3 (1,464)	24.0 (276)	232.8 (670)	497.4 (1,916)	368.6 (3,408)	343.9	354.9
Females									
35–44	0.4 (1)	8.6 (16)	4.2 (19)	23.8 (46)	56.7 (110)	80.1 (134)	72.9 (288)	65	62
45–54	2.8 (6)	25.4 (43)	15.3 (65)	45.6 (46)	133.2 (141)	197.2 (179)	172.1 (377)	175	177
55–64	10.9 (24)	51.9 (46)	25.4 (85)	85.9 (45)	268.5 (166)	412.6 (141)	335.7 (360)	330	339
65–74	18.7 (37)	103.7 (31)	34.3 (82)	121.2 (28)	576.9 (182)	877.1 (66)	627.1 (274)	565	555
75–84	48.4 (55)	143.9 (9)	55.5 (68)	143.0 (19)	921.8 (190)	1,303.9 (21)	999.2 (229)	868	904
Total: 35–84[c]	7.5 (123)	39.4 (145)	17.5 (319)	59.4 (184)	220.9 (789)	328.8 (541)	260.6 (1,528)	244.5	246.2

[a] Unweighted No. of sample deaths are given *in parentheses*.
[b] Death rates are based on 1966–68 NMS and August 1967 CPS. In column 5, values are for prostate cancer in males and breast cancer in females.
[c] Rates are age adjusted by the direct method to the 1940 U.S. population 35–84 yr old.

TABLE 8.—*The 1966–68 weighted No. of observed (Obs) and expected (Exp) deaths and SMR by major cancer site for a representative sample of U.S. whites 35–84 years old who never smoked cigarettes*

Primary site	ICD (Seventh Revision) code No.	9,525,416 male nonsmokers				28,945,730 female nonsmokers			
		Obs	Sample No.[a]	Exp[b]	SMR[c]	Obs	Sample No.[a]	Exp[b]	SMR[c]
Total malignant neoplasms	140–205	106,478	670	157,798	67	237,261	789	271,588	87
Buccal cavity and pharynx	140–148	2,728	56	4,403	62	1,702	28	3,248	52
Digestive organs and peritoneum	150–159	45,029	135	48,457	93	74,968	180	87,580	86
Esophagus	150	1,782	37	3,216	55	1,601	26	1,992	80
Stomach	151	7,920	17	10,651	74	10,553	27	13,325	79
Colon	153	17,967	44	13,553	133	30,438	61	32,446	94
Rectum	154	7,560	18	6,463	117	8,117	18	10,573	77
Liver, gallbladder, and biliary passages	155–156	3,282	6	5,272	62	11,717	22	13,311	88
Pancreas	157	4,759	9	7,870	60	11,658	23	13,983	83
Respiratory system	160–165	18,038	115	46,581	28	12,908	131	21,711	59
Lung, bronchus, and trachea	162–163	8,733	108	39,991	22	9,079	123	17,122	53
Breast	170					55,051	184	58,659	94
Female genital system	171–176					38,385	124	41,576	92
Cervix	171					7,790	30	11,642	67
Corpus uteri	172					1,126	6	1,937	58
Uterus, NOS[d]	173–174					5,348	16	6,316	85
Ovary, fallopian tube, and broad ligament	175					21,563	68	19,393	111
Male genital system	177–179	13,784	278	16,563	83				
Prostate gland	177	13,534	276	16,263	83				
Urinary system	180–181	5,881	18	9,578	61	6,816	12	7,926	86
Nervous system	193	2,478	10	2,609	95	6,830	24	5,543	123
Lymphomas	200–202	4,187	14	6,417	65	7,991	21	9,131	88
Leukemia	204	3,361	12	5,238	64	8,754	24	10,107	87
Total malignant neoplasms[e]	140–205	106,478	670	152,037[f]	70	237,261	789	250,521[f]	95

[a] Sample No. of deaths from which weighted No. of observed deaths is calculated with the use of poststratified ratio estimation procedure (*see* text).
[b] Expected No. of deaths from cancer is based on 5-yr, age-specific death rates determined for U.S. whites (1966–68) from the t. 1966–68 NMS and August 1967 CPS (*see* text).
[c] SMR is the percentage defined as Obs÷Exp × 100.
[d] NOS = not otherwise specified.
[e] Expected No. of deaths from cancer is based on 5-yr, age-specific death rates for U.S. whites in 1970 (*8*).

male and 95% for female nonsmokers. For male nonsmokers, the lowest SMR values occurred for the traditional smoking-related cancer sites of the respiratory system, buccal cavity and pharynx, esophagus, pancreas, and urinary system, with a combined SMR of 89%, whereas the remaining cancer sites had a combined SMR of 91%. However, the lowest SMR values for female nonsmokers occurred for the cancer sites of the lung and buccal cavity and pharynx, with a combined SMR of 53%; the SMR values for most of the other sites were between 80 and 95%. Because of the complex sampling and weighting procedures used in these data, confidence limits on an individual SMR cannot be precisely calculated, although approximate limits can be determined with the use of the sample number of deaths, if one assumes they are subject to a Poisson distribution (*2*).

DISCUSSION

The 1966-68 NMS represents the first and only attempt to date to determine death rates for all causes of death among a representative sample of the U.S. population classified according to smoking status. The basic findings in this paper show that during 1966-68, U.S. whites 35-84 years old who never smoked cigarettes have age-adjusted death rates from cancer and all causes substantially below those of the population as a whole and only about one-half the corresponding rates of those who currently smoke cigarettes. Phrased in other terms, whites who never smoked cigarettes have overall SMR which are substantially less than 100%. Specific comparisons vary by a few percentage points with different reference populations or with direct versus indirect standardization, but the basic patterns remain the same. Furthermore, these results indicate that most of the reduction in cancer deaths among male nonsmokers occurs among the traditional smoking-related cancer sites of the respiratory system, buccal cavity and pharynx, esophagus, pancreas, and urinary system. The single greatest reduction in cancer deaths among female nonsmokers occurs in the respiratory system; however, small reductions also occur at many other sites.

These results provide a methodologically sound data set on which to base comparisons with other popula.

tions at low risk to cancer. In other words, any nonsmoking population who has death rates similar to the 1966-68 NMS nonsmokers is probably at low risk solely because they do not smoke. These results help delineate the specific influence of cigarette smoking on mortality rates from other health-related factors.

REFERENCES

(1) Advisory Committee to the Surgeon General of the Public Health Service. Smoking and health. Washington, D.C.: U.S. Govt Print Off, 1964 [DHEW publication No. (PHS)1103].

(2) HAENSZEL W, LOVELAND DB, SIRKEN MG. Lung-cancer mortality as related to residence and smoking histories. I. White males. J Natl Cancer Inst 1962;28:947-1101.

(3) HAENSZEL W, TAEUBER KE. Lung-cancer mortality as related to residence and smoking histories. II. White females. J Natl Cancer Inst 1964;32:803-838.

(4) GODLEY F, KRUEGEL DL. Cigarette smoking and differential mortality. New estimates from representative national samples. In: Health: United States, 1975. Rockville, Md.: National Center for Health Statistics, 1976:463 [DHEW publication No. (HRA)76-1232].

(5) ENSTROM JE. Cancer and total mortality among active Mormons. Cancer 1978;42:1943-1951.

(6) ———. Rising lung cancer mortality among nonsmokers. JNCI 1979;62:755-760.

(7) GROVE RD, HETZEL AM. Vital statistics rates in the United States, 1940-1960. Washington, D.C.: U.S. Govt Print Off, 1968 [DHEW publication No. (PHS)1677].

(8) National Center for Health Statistics. Vital statistics of the United States, annual reports for 1966, 1967, 1968, and 1970. Rockville, Md.: National Center for Health Statistics, 1974.

(9) ———. Cigarette smoking status—June 1966, August 1967, and August 1968. Monthly vital statistics report. Vol 18, No. 9 (suppl, December 5, 1969). Rockville, Md.: National Center for Health Statistics, 1969.

(10) HAMMOND EC. Smoking in relation to the death rates of one million men and women. Natl Cancer Inst Monogr 1966;19: 127-204.

(11) KAHN HA. The Dorn study of smoking and mortality among U.S. veterans: Report on eight and one-half years of observation. Natl Cancer Inst Monogr 1966;19:1-125.

(12) U.S. Bureau of the Census. Estimates of the population of the United States by age, color, and sex, July 1, 1967. Current population reports, series P-25, No. 385. Washington, D.C.: U.S. Govt Print Off, 1967.

(13) KITAGAWA EM, HAUSER PM. Differential mortality in the United States: A study of socioeconomic epidemiology. Cambridge: Harvard Univ Press, 1973.

(14) FRANKEL MR. Inference from survey samples: An empirical investigation. Ann Arbor, Mich.: Institute for Social Research, Univ Michigan, 1971.

(15) KIRCHOFF H, RIGDON RH. Smoking habits of 21,612 individuals in Texas. J Natl Cancer Inst 1956;16:1287-1304.

APPENDIX

National Mortality Survey

The NMS is based on a systematic random sample of all deaths occurring and registered in the United States.[5] Death registration is essentially 100% complete and is performed on a state-of-occurrence basis with the use of the official State death certificate. The States generally follow the federally recommended model death certificate. Date, place, and cause of death, age, sex, and other demographic characteristics are important legal facts noted on the document. One can extend the research potentials of death certificates by using them as a sampling frame for surveys (such as the NMS) in which additional information is collected through informants, funeral directors, physicians, and others named on the certificate.

The scope of the NMS is all deaths of persons 35-84 years old that occurred in the United States during 1966, 1967, or 1968. The sample was stratified for month of death, geographic area, age, and selected causes of death related to smoking. The total sample of 19,526 decedents between 35 and 84 years of age was drawn in two stages. 1) A 10% sample is represented in the monthly "Current Mortality Survey" of death certificates submitted to the NCHS on microfilm by all States, the District of Columbia, and other death registration areas. 2) The strata of the 10% sample are subsampled; each 26th case was taken on the average. The overall sampling rate is therefore 1/260. The demographic and medical information coded from the sampled death certificate is supplemented by information from a mailed questionnaire on smoking habits, marital status, income, and other characteristics of the deceased. One-half of the respondents are spouses or ex-spouses of the deceased, one-fourth are either parents or offspring, one-eighth are siblings, and the remaining 10% are comprised of more distant relatives, friends, neighbors, and other associates. The overall response rate of 95% in the NMS is comparable to that of the CPS.

After follow-up queries of questionnaires with incomplete, inconsistent, or otherwise inadequate responses, smoking information was still missing for about 8% of the sample deaths. For these cases, smoking history was assigned by computer processing of the records of the decedents. This procedure, in the absence of better information, assumes that the characteristics for which reported information is missing are like those of demographically similar individuals for whom the information in question was reported.

This processing was performed by the NCHS, which also supplied the magnetic tape of 1966-68 NMS records. Each record combines the smoking and other NMS questionnaire information with that from the death certificate. In addition, each record contains an inflation factor (sample weight) that reflects the sampling fraction applicable to the decendent's demographic cause-of-death stratum. This sample weight is determined by a poststratified ratio estimation procedure which forces the total number of deaths estimated from sampling to agree with the total number of deaths from which the sampling was done.[5] The weight is a function of age, year of death, and cause of death and varies from about 10 to 640. Appendix table 1 shows an example of how the total weighted number of 55,051 deaths is calculated from the 184 sample deaths due to breast cancer among U.S. white females who never smoked cigarettes. The age-sex distributions of the total weighted deaths from all causes according to smoking status are given in Appendix table 2. The underlying cause of death (item #18 of the death

APPENDIX TABLE 1.—*Sample calculation of 1966–68 average annual age-specific and age-adjusted death rates for breast cancer for U.S. white females who never smoked cigarettes by the poststratified ratio estimation procedure*

Age, yr	Unweighted No. of sample deaths by yr				Poststratified ratio estimation sample weight by yr			Total weighted No. of sample deaths, 1966–68[a]	August 1967 population-at-risk	Average annual death rate (deaths/100,000), 1966–68	Proportional distribution of 1940 U.S. population
	1966	1967	1968	1966–68	1966	1967	1968				
35–44	18	15	13	46	81	82	84	3,780	5,301,563	23.8	0.345440
45–54	17	12	17	46	163	162	166	7,537	5,514,745	45.6	0.292283
55–64	17	10	18	45	313	319	314	14,163	5,496,482	85.9	0.199205
65–74	11	7	10	28	639	640	611	17,619	4,847,416	121.2	0.120142
75–84	5	7	7	19	639	640	611	11,952	2,785,516,	143.0	0.042930
Total: 35–84	68	51	65	184				55,051	23,945,722	59.4[b]	1.000000

[a] Weight No. of sample deaths in each 10-yr age group is obtained by multiplication: Unweighted No. of sample deaths in each death year times the poststratified ratio estimation sample weight for the same death year; add results for the 8-yr total.[5]
[b] Age-adjusted death rate is calculated by the direct method with the 1940 U.S. population 35–84 yr of age as the standard.

certificate) is determined from information provided by physicians and coded at the NCHS according to the rules of the Seventh Revision of the ICD (*7, 8*).

Current Population Survey

The CPS (*9*) is the source of estimates of the nonsmoking and smoking populations. Smoking information was collected by the U.S. Bureau of the Census in its CPS of August 1967. Designed to be a representative sample of the Nation's noninstitutionalized population, the total consisted of about 50,000 households (residents of military bases, homes for the sick and aged, and other institutions were excluded). The overall response rate was 95%; of the 95,532 adults in the sample, those between the ages of 35 and 84 years numbered 60,920 and included 25,266 white males and 29,308 white females. Based on this sample, the age-sex distributions of the noninstitutionalized population according to smoking status are shown in Appendix table 3. Also shown for comparison is the CPS for July 1, 1967 (*12*), which represents the total resident population.

The U.S. Bureau of the Census interviewers recorded information about household and personal characteristics of those in the sample. Demographic and socioeconomic information for all persons in the household was reported by a single adult respondent. Smoking information was obtained from each adult responding for himself, if this was possible during the one interview. However, the smoking information for [illegible] not at home was also obtained from the person interviewed for the household. Inasmuch as men are more likely than women to be away from home when the interviews take place, only about one-third of the men, as compared with three-fourths of the women in the sample, reported on their smoking habits. Smoking information was missing for only about 3% of the sample.

Estimates of population frequencies of nonsmokers by demographic characteristics were obtained from

APPENDIX TABLE 2.—*Estimated 1966–68 U.S. total deaths of whites based on representative sample of deaths from the NMS by age, sex, and cigarette smoking status*

Age, yr	U.S. whites who never smoked cigarettes[a]		U.S. whites who currently smoke cigarettes[a]		All U.S. whites[a]	
	Male	Female	Male	Female	Male	Female
35–39	6,700 (81)	7,840 (98)	28,181 (346)	12,067 (148)	38,858 (476)	22,600 (279)
40–44	8,521 (105)	16,285 (200)	49,359 (608)	21,116 (262)	66,804 (824)	42,220 (521)
45–49	13,734 (93)	20,697 (132)	74,458 (528)	33,045 (221)	103,119 (732)	60,191 (398)
50–54	25,868 (170)	34,463 (222)	118,466 (850)	39,771 (273)	169,694 (1,207)	84,686 (566)
55–59	39,723 (172)	47,388 (179)	146,149 (774)	48,336 (206)	235,209 (1,193)	110,514 (448)
60–64	58,025 (265)	79,733 (309)	168,468 (909)	46,688 (209)	293,031 (1,527)	141,862 (589)
65–69	90,047 (273)	131,184 (304)	153,307 (691)	48,520 (139)	327,846 (1,351)	195,314 (488)
70–74	109,536 (397)	217,599 (515)	144,226 (656)	41,403 (113)	363,764 (1,547)	280,660 (698)
75–79	158,226 (498)	290,082 (708)	113,174 (482)	39,244 (96)	377,154 (1,426)	349,985 (862)
80–84	157,309 (478)	298,569 (691)	70,960 (270)	20,659 (53)	309,449 (1,035)	337,428 (787)
Total: 35–84	667,689 (2,532)	1,143,840 (3,358)	1,066,748 (6,114)	350,849 (1,720)	2,284,928 (11,318)	1,625,460 (5,636)

[a] The *first No. in each column* is the weighted No. of deaths with the use of weight based on the poststratified ratio estimation procedure.[5] *Numbers in parentheses* are the unweighted No. of sample deaths used in the NMS.[5]

APPENDIX TABLE 3.—*Estimated 1967 age distribution of U.S. whites 35–84 years old by cigarette smoking status based on the CPS*

Age, yr	U.S. whites who never smoked cigarettes		U.S. whites who currently smoke cigarettes		All U.S. whites			
	Male[a]	Female[a]	Male[a]	Female[a]	Male (August 1967)[a]	Male (July 1, 1967)[b]	Female (August 1967)[a]	Female (July 1, 1967)[b]
35–39	1,168,421	2,527,542	2,779,025	2,205,703	4,779,120	5,002,000	5,137,756	5,172,000
40–44	1,185,151	2,774,021	3,064,389	2,307,515	5,235,199	5,378,000	5,588,969	5,627,000
45–49	1,195,262	2,788,674	2,894,887	2,248,005	5,054,911	5,161,000	5,434,324	5,464,000
50–54	1,119,008	2,726,071	2,541,111	1,819,272	4,647,741	4,723,000	4,965,615	5,001,000
55–59	1,009,872	2,778,288	2,084,528	1,350,491	4,095,290	4,153,000	4,454,802	4,489,000
60–64	923,158	2,718,194	1,559,549	843,503	3,392,068	3,452,000	3,835,309	3,875,000
65–69	937,471	2,564,226	956,542	484,001	2,663,209	2,714,000	3,221,374	3,269,000
70–74	883,178	2,283,190	552,483	241,868	2,009,687	2,067,000	2,650,990	2,726,000
75–79	695,297	1,834,020	326,279	107,273	1,439,539	1,462,000	2,002,833	2,038,000
80–84	408,598	951,496	114,336	44,890	689,562	800,000	1,012,735	1,195,000
Total: 35–84	9,525,416	23,945,722	16,873,129	11,652,521	34,006,326	34,912,000	38,304,207	38,856,000

[a] Numbers are from the August 1967 CPS of noninstitutionalized U.S. population by cigarette smoking status (*9*).
[b] Numbers are from the July 1, 1967 CPS of total resident U.S. population (*12*).

inflation factors supplied by the U.S. Bureau of the Census in a magnetic tape of the sample person records. The inflation factors take into account the sampling fraction and demographic characteristics of those in noninterviewed sample households. Missing demographic characteristics of those in interviewed households were computer assigned except for smoking information. "Not stated" or otherwise unknown smoking status is distributed statistically in proportion to the relative frequencies in the known smoking categories. This procedure is followed, for lack of evidence of a smoking-related pattern of nonresponse, so the best estimate of population can be provided. Because it incorporates U.S. census control totals, the CPS is representative of geographic units, age, sex, and other demographic characteristics of the national population.

Standard Death Rates

Expected numbers of deaths in table 3 are the product of a set of standard death rates for 1966–68 (Appendix table 2) and the August 1967 CPS estimates of whites who never smoked cigarettes (Appendix table 3). The standard rates reflect the mortality experience of the total white population (irrespective of smoking history). These rates are specific for 5-year age groups, sex, and the underlying causes of cancer death shown in table 3. We computed the standard rates by dividing the total of the NMS death estimates for 1966, 1967, and 1968 by the CPS population estimate for August 1967. Thus both expected and observed deaths were derived from the same data base of deaths. In other words, for each cause of death, the total number of expected deaths equals the total number of observed deaths for all white males and females. This procedure minimizes some biases that otherwise may be present in the SMR when the expected deaths are based on an arbitrary standard such as the 1970 U.S. white population.

Sources of Error

Various types of systematic errors likely to affect the estimates of mortality from smoking are discussed below. Also, the reliability of the estimated statistics in relation to sampling error and other sources of random variability is evaluated.

1) Matching between deaths and populations: The major source of bias affecting the results of this study arises from the cross-sectional two-sample design. Theoretically, the decedents are a subset of the total study population; ideally, the prospective design best approximates this concept. The American Cancer Society and U.S. Veterans cohorts (*10, 11*) are prospective studies in which a beginning sample is initially questioned regarding current and past smoking habits. The cohort is followed longitudinally, and its decedents are recorded with respect to length of time in the study. Death rates are computed as follows: the total number of deaths divided by the total number of person-years lived during the study period by persons in various subgroups. These persons are classified by smoking habits and other characteristics. The decedents of a prospective study, included in the original sample, match exactly with an individual in the study population.

The present cross-sectional study, however, must rely on statistical matching between the two independently drawn samples: decedents and population. This problem is common to conventional vital statistics rates (*7, 8*) and must be balanced with the expense, time, and follow-up problems of prospective studies.

Because mortality risk is measured by ratio of deaths to population, the matching problem is a basic consideration in the evaluation of errors from sampling, the data-collection process, and measurement techniques.

2) The samples: A comparability problem exists in the extent to which the target populations differentially represent their universe. The universe of all

deaths occurring to persons 35-84 years old in the United States during 1966-68 is represented by official registrations of deaths occurring in the United States in any of the 3 years, irrespective of month of death or usual residence of the deceased. However, the population interviewed in the CPS consists of individuals who were not residents of institutions in the United States as of August 1967; this count was at about the midpoint and is believed to represent the 1966-68 population adequately. The CPS practice of replacing part of the sample every 4 months (rather than monthly) extends parts of the target population to cover a span of 7 months, May to November 1967.

Conversely, exclusion of residents of institutions introduces a systematic bias that understates the population and thereby inflates death rates. A first consideration is also to exclude deaths of residents of institutions. The rationale for not doing so is as follows: *a*) In terms of death rates, the bias is really a function of person-years of excluded institutional population. Fragmentary evidence suggests that many persons in the institutional population are there for only a fraction of a year.[4] *b*) There is no sound basis for the estimation of the person-years of excluded institutionalized population for the various socioeconomic groups classified by smoking status. *c*) The net effects of the bias are likely to be canceled in measures of differential mortality like the SMR. *d*) The bias is most likely to affect only the extremely old, a small fraction of the study population for whom differential mortality is greatly reduced and not of major interest. Only for persons 80-84 years of age is the excluded population more than 5% of the total (Appendix table 3).

3) The respondents: Unlike sizable proportions of the population sample who could provide information on their smoking habits and other personal characteristics, no such data were available from the death sample. Moreover, a mailed questionnaire was used in the death survey, whereas a household interview was conducted for collection of data on the living population. Other sources report substantial agreement between the smoking habits reported for a decedent retrospectively and his reports as obtained in a household interview prior to death (*2*).[4] With regard to social and demographic characteristics, evidence indicates agreement between the 1960 NMS responses and the matched 1960 census reports (*13*).

The overall response rates to the NMS and the CPS were nearly equal at 95%; these high levels eliminate any serious bias from differential nonresponse. There is no evidence on nonresponse differentials by smoking status from this investigation. In retrospective studies, in which nonresponse may be as much as 32% overall and even more among smokers, the bias in the mortality ratio of the respondent has been estimated to be about a 0.2 or 0.3 overstatement because of lower mortality ratios of the nonrespondents (*1*). Again, the low level of nonresponse in the present study (5%) indicates that nonresponse is negligible as a source of bias.

4) Measurement of cigarette smoking: The operational definitions of lifetime history concepts of never smoking cigarettes are almost exactly the same in death survey (<5 packs of cigarettes) and population survey (<100 cigarettes). Copies of the actual NMS and CPS questionnaires are shown elsewhere.[4] Although the slight variations in these definitions do not appear to be a problem, a comparability problem does arise from the fact that those reporting for the decedent tend to understate cigarette consumption compared with what the decedent would have reported as his habits. Haenszel et al. (*2*) noted that this bias is largest with respect to medium and heavy smokers but virtually nonexistent for nonsmokers.

5) Sampling and data processing: Complex sampling design, coding of reported information, construction of data files, and production of the final data tables are all possible sources of systematic error. The assumption is that these errors are negligible. Table 1 and Appendix table 3 show that the study estimates of death rates and population-at-risk are representative of the United States for the study period. Table 1 illustrates the good agreement between the total age-specific death rates as calculated in the survey and the 1966-68 U.S. vital statistics (*8*). The good agreement between the total age-specific U.S. white population as determined by the August 1967 and the July 1, 1967, CPS (*12*), which was used to calculate the 1967 U.S. death rates (*8*), is depicted in Appendix table 3.

6) Reliability: In addition to systematic error, the estimates reported in this study contain some random error. This kind of variability is inherent in the sampling of deaths and population at a particular time. In more detailed analyses, random variation may be present in the indicators of mortality that are used for empirical tests of hypotheses. The 1964 Surgeon General's Report evaluated the stability of the SMR by a main assumption that deaths occurring within a given period are generated by a Poisson process (*1*). The random variability of SMR was small and depended on the number of deaths. The analysis was of 26,000 deaths in 7 prospective studies. Inasmuch as this study is based on about 20,000 sample deaths, one can assume that variability in its SMR is a negligible problem for the major causes of death.

7) Sampling error: In addition to variability over time, the SMR may contain some random variation due to sampling. Both death and population estimates are based on complex, multistage sample designs. Another analysis of the 1967 CPS shows that simple random sampling standard errors understate the true standard errors of estimated population statistics (*14*). Standard tests of statistical significance are therefore inapplicable to this study.

Hammond (*10*) used a frequency of 10 deaths as a minimum for standards of reliability. Although arbitrary, the same rule of thumb (adopted in this study) primarily affects deaths when classified by detailed underlying causes. The sample numbers of deaths from the 1966-68 NMS are presented in tables 1-3.

8) Response error: At the level of the individual, reported information may vary from one observation to another. This variation may introduce error in estimates of fixed characteristics like a history of ever smoking. Kirchoff and Rigdon (*15*) compared the smoking habits of a group of Texans and reported different smoking habits for the same reference period when respondents were reinterviewed. This difference partly reflects some memory lapse and some response error in smoking information. We assume that the extent to which this source of variation affects estimates of the present study is negligible but unknown.