

# Misclassification Rates for Current Smokers Misclassified as Nonsmokers

A. Judson Wells, PhD, Paul B. English, PhD, MPH, Samuel F. Posner, PhD, Lynne E. Wagenknecht, DrPH, and Eliseo J. Perez-Stable, MD

## ABSTRACT

**Objectives.** This paper provides misclassification rates for current cigarette smokers who report themselves as nonsmokers. Such rates are important in determining smoker misclassification bias in the estimation of relative risks in passive smoking studies.

**Methods.** True smoking status, either occasional or regular, was determined for individual current smokers in 3 existing studies of nonsmokers by inspecting the cotinine levels of body fluids. The new data, combined with an approximately equal amount in the 1992 Environmental Protection Agency (EPA) report on passive smoking and lung cancer, yielded misclassification rates that not only had lower standard errors but also were stratified by sex and US minority/majority status.

**Results.** The misclassification rates for the important category of female smokers misclassified as never smokers were, respectively, 0.8%, 6.0%, 2.8%, and 15.3% for majority regular, majority occasional, US minority regular, and US minority occasional smokers. Misclassification rates for males were mostly somewhat higher.

**Conclusions.** The new information supports EPA's conclusion that smoker misclassification bias is small. Also, investigators are advised to pay attention to minority/majority status of cohorts when correcting for smoker misclassification bias. (*Am J Public Health*. 1998;88:1503-1509)

In recent years, concerns have been raised that the association between passive smoking and lung cancer could in whole or in part be ascribed to an upward bias caused by misclassification of smokers as never smokers.<sup>1</sup> These concerns arise because a small proportion of subjects is known who report as never having smoked, through either tests of their body fluids or repeated questionnaires, to probably smoke currently or to have smoked in the past.

Because most of the passive smoking studies involve spousal exposure of women who say they have never smoked, such misclassification could create the impression of a passive smoking effect because smokers tend to marry smokers. Therefore, if misclassification is appreciable, more of the smokers misclassified as never smokers will be found among those reporting as never smokers married to smokers than among those reporting as never smokers married to nonsmokers. In the case of lung cancer, the high relative risks of even a small percentage of misclassified smokers would increase the incidence of cancer among the self-reported never smokers married to the smokers relative to those married to the nonsmokers and would create a spurious passive smoking effect. The emphasis here is on the misclassification of self-reported never smokers rather than a category that would include former smokers, because the various passive smoking studies deal almost entirely with subjects who never smoked.

Several methods<sup>1-4</sup> have been proposed to determine the effect of smoker misclassification on the observed passive smoking relative risks. These methods require a knowledge of misclassification rates for current and former smokers, the degree of concordance of smoking status between spouses, the prevalence of smoking in the subject population, and the proportion of never smokers who are exposed to tobacco smoke. All of these factors are important, but

among the most important, and least easily derived from the published literature, are the degree to which current smokers are misclassified as never smokers and the level at which they smoked.

Misclassification rates for current smokers can be measured with biomarkers in body fluids. The best biomarker is thought to be cotinine,<sup>5</sup> a metabolite of nicotine but with a longer half-life of about 20 hours, which is specific to tobacco smoke and can be measured in saliva, urine, or serum. Unfortunately, none of the passive smoking studies includes detailed information on cotinine levels in the body fluids of subjects. Therefore, estimates of misclassification rates for smokers must be derived from other studies.

In most of the literature on cotinine levels, one or more of the essential factors needed for the passive smoking bias calculation is missing. Such studies lack either breakdown by sex or separation of self-reported never smokers from self-reported former smokers, or the misclassification rate is stated simply as the number or percent of reported nonsmokers (never smokers plus former smokers) who had cotinine levels

Paul B. English is with Impact Assessment, Inc, a contractor for California Department of Health Services, Environmental Health Investigations Branch, Emeryville, Calif. Samuel F. Posner is with the Medical Effectiveness Research Center for Diverse Populations, Division of General Internal Medicine, Department of Medicine, University of California, San Francisco. Lynne E. Wagenknecht is with the Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC. Eliseo J. Perez-Stable is with the Medical Effectiveness Research Center for Diverse Populations, Division of General Internal Medicine, Department of Medicine, University of California, San Francisco.

Requests for reprints should be sent to A. Judson Wells, PhD, 5 Ingleton Circle, Kennett Square, PA 19348.

This paper was accepted March 11, 1998.

above a certain cutpoint that was thought to distinguish smokers from nonsmokers.<sup>6-12</sup> Such a number is not useful in determining the smoker misclassification bias by any of the more accurate methods because these assume that the risk (e.g., for lung cancer) is proportional to the cotinine level and, therefore, require individual cotinine levels to determine the misclassified smokers' risk. Furthermore, the misclassification rates must be based on the number of smokers in the cohort, not the number of nonsmokers, because the more smokers there are in a given cohort, the greater is the probability of finding misclassified smokers.

Some data on current smoker misclassification rates appear in the 1992 report on lung cancer and passive smoking by the US Environmental Protection Agency (EPA).<sup>4</sup> When the EPA report was written, accurate misclassification rates for current smokers were scarce, so the EPA used all sources (7) in which either cotinine levels of misclassified individuals or fairly detailed ranges of same were available. For 4 of its 7 sources, the EPA relied on personal communications with the authors to obtain the required data. The purpose of this paper is to provide additional data, beyond that in the EPA report, on the misclassification rates for current smokers who reported as never smokers in order to allow greater accuracy in calculating the bias in passive smoking relative risks caused by smoker misclassification.

## Methods

Three of the authors of this paper were also lead authors of 3 studies in which detailed smoking status questions were asked.<sup>10-12</sup> These authors, via personal communications, have provided data from these studies on cotinine levels of individuals who reported as either never smokers, former

smokers, or current smokers. In addition, the raw data from the 7 studies in the EPA report<sup>4</sup> have been reexamined to correct one error and to include, from 4 of the studies, data on males that were not available in the EPA report. These sources provide the new information given here on which to base current smoker misclassification rates.

Definitions of never smokers and former smokers vary among the studies. Most frequently, never smokers are defined as those subjects who have smoked fewer than 100 cigarettes during their lifetimes. Former smokers are usually defined as those who have refrained from smoking for a year or more. Nonsmokers as used here means never smokers plus former smokers.

Subjects from a given study who reported as never or former smokers but who had cotinine levels greater than 10% of the mean cotinine level for self-reported smokers from that study were considered misclassified current smokers. Misclassified smokers from a given study with cotinine levels greater than 30% of the self-reported smoker mean for that study are defined as "regular" smokers because it is assumed that they have a lung cancer risk equal to the self-reported current smokers from that study. This assumption is supported by the finding that, for the 7 EPA studies<sup>4</sup> and the 3 new studies reported here, the mean cotinine level for the misclassified regular smokers equals 98% of the mean cotinine level for all of the self-reported current smokers.

The rest of the misclassified smokers, those with cotinine levels between 10% and 30% of the self-reported smoker mean cotinine level for a given study, are defined as "occasional" smokers. This would be equivalent to smoking about 1 to 4 cigarettes per day.<sup>13</sup> Based on the 10 EPA and new studies, these misclassified occasional smokers have a mean cotinine level that is about 18% of the self-reported current smokers' mean

level. Therefore, assuming a linear relationship with lung cancer risk, occasional smokers would have about 18% of the risk of the self-reported smokers. The cutpoint of 10% of the mean cotinine level of the self-reported smokers used to distinguish misclassified occasional smokers from true nonsmokers is, of course, arbitrary. Certainly some misclassified smokers have cotinine levels below this point. However, confusion between them and highly exposed passive smokers would increase as this cutpoint decreases. A lower cutpoint would result in higher misclassification rates for the occasional smokers, but their effect would be offset by a lower disease risk for the occasional category. The net effect on the smoker misclassification bias calculation is minimal.<sup>4</sup> The advantage of the 10% cutpoint is that it avoids inclusion of true passive smokers.

During the course of the investigation, an appreciable difference was seen between the misclassification rates for US Blacks and Latinos and the rates for Whites in the United States and majority groups in various other countries. Therefore, the misclassification data for both the 3 new studies and the 7 EPA studies are stratified by (1) sex, (2) US minority/majority status, (3) self-reports of never smokers or former smokers, and (4) determination of whether the misclassified smokers were considered occasional smokers or regular smokers.

Misclassification rates for a given category are calculated by dividing the total number misclassified for all studies in that category by the total number of self-reported current smokers in that category. For example, the number of misclassified majority female regular smokers would be divided by the number of self-reported majority female regular smokers. This type of rate is used as input for the Wells-Stewart method,<sup>4</sup> which the EPA selected.<sup>4</sup> So-called true rates, which include the misclassified smokers in

**TABLE 1—New Studies in This Report of Cotinine Levels of Nonsmokers and Smokers**

Study	Cohort Studied	Cotinine Source	Mean Cotinine Level, ng/mL, of Self-Reported Smokers
English et al. <sup>10</sup>	3343 pregnant White and Black females from Child Health and Development Studies, Oakland, Calif, enrolled between 1959 and 1966	Serum	Whites, 137 Blacks, 148
Perez-Stable et al. <sup>12,15</sup>	189 Mexican Americans aged 20–74 who were part of HHANES from Arizona, California, Colorado, New Mexico, and Texas, enrolled 1982–1983	Serum	Females, 144 Males, 161
Wagenknecht et al. <sup>11,16</sup>	4973 Black and White subjects aged 18–30 from the CARDIA study from Alabama, California, Illinois, and Minnesota, enrolled 1985–1986	Serum	White females, 176 Black females, 251 White males, 210 Black males, 245

Note. HHANES = Hispanic Health and Nutrition Examination Survey; CARDIA = Coronary Artery Risk Development in Young Adults.

**TABLE 2—Numbers of Misclassified Current Smokers From the 3 New Studies in This Report, Stratified by Smoking Level, Sex, and Majority/US Minority Status<sup>a</sup>**

Study	Smoking Status of Misclassified Current Smokers	Self-Reported Smoking Status, No.			Cotinine Level Ranges, ng/mL, for Misclassified Current Smokers <sup>b</sup>
		Never	Former	Current	
<b>US non-Hispanic White females</b>					
English et al. <sup>10</sup>	Occasional	6	8	134	14–41
	Regular	5	3	695	>41
	Total subjects	1136	308	829	
Wagenknecht et al. <sup>11</sup>	Occasional	4	9	71	18–53
	Regular	1	4	282	>53
	Total subjects	668	250	353	
Totals, US non-Hispanic White females	Occasional	10	17	205	
	Regular	6	7	977	
	Total subjects	1804	558	1182	
<b>US minority females</b>					
English et al. <sup>10</sup> (Black)	Occasional	10	0	53	15–44
	Regular	8	2	321	>44
	Total subjects	641	55	374	
Perez-Stable et al. <sup>12</sup> (Mexican American)	Occasional	2	1	7	14–43
	Regular	0	0	20	>43
	Total subjects	77	20	27	
Wagenknecht et al. <sup>11</sup> (Black)	Occasional	8	6	68	25–75
	Regular	11	10	386	>75
	Total subjects	869	124	454	
Totals, US minority females	Occasional	20	7	128	
	Regular	19	12	721	
	Total subjects	1587	199	855	
<b>US non-Hispanic White males</b>					
Wagenknecht et al. <sup>11</sup>	Occasional	3	5	55	21–63
	Regular	6	8	251	>63
	Total subjects	654	178	306	
<b>US minority males</b>					
Perez-Stable et al. <sup>12</sup> (Mexican American)	Occasional	1	3	17	16–48
	Regular	2	3	49	>48
	Total subjects	47	45	66	
Wagenknecht et al. <sup>11</sup> (Black)	Occasional	8	7	45	24–73
	Regular	14	6	370	>73
	Total subjects	599	103	415	
Totals, US minority males	Occasional	9	10	62	
	Regular	16	9	419	
	Total subjects	646	148	481	

<sup>a</sup>The numbers of self-reported never, former, and current smokers are from personal communications from Drs English, Perez-Stable, and Wagenknecht. The numbers of self-reported current occasional smokers include all subjects who report themselves as current smokers but whose cotinine level is <30% of the mean cotinine level for that category as shown in Table 1. This is consistent with the practice in the EPA report.<sup>4</sup>

<sup>b</sup>The cotinine level ranges for misclassified current smokers are also derived from the mean levels for self-reported current smokers shown in Table 1, 10% to 30% for misclassified occasional smokers and >30% for misclassified regular smokers. The "smoking status" of the misclassified smokers is based on their cotinine level.

the denominator with the self-reported smokers, are used in another method,<sup>1</sup> but their use results in complicated iterative calculations with no increase in the accuracy of the bias calculation.<sup>4</sup>

Standard errors were calculated as follows:  $SE = (m(1 - m)/N)^{1/2}$ , where  $m$  is the misclassification rate, and  $N$  is the total number of current smokers for a particular category.<sup>14</sup> The combined estimated misclassification rates and standard errors for smokers who reported as never smokers or as former smokers are computed by cotinine-determined

smoking status for all available data. Combined misclassification rates were calculated for the 3 new studies, for the 7 EPA studies, and for all 10 studies.

## Results

The 3 new studies in this report are described briefly in Table 1. The numbers of subjects, the body fluid that is the cotinine source, and the mean cotinine levels of the self-reported current smokers are shown. The

numbers of smokers misclassified as never smokers and former smokers for the 3 studies are shown in Table 2 along with the numbers of self-reported current smokers. The combined misclassification rates and standard errors for smokers of various categories who report as never smokers or former smokers are shown in Table 3. Similar information for the 7 studies in the EPA report<sup>4</sup> are shown in Tables 4, 5, and 6. The combined misclassification rates for all 10 studies are shown in Table 7. The misclassification rates for the new studies in Table 3 are reasonably consis-

**TABLE 3—Misclassification Rates and Standard Errors for Current Smokers Self-Reported as Either Never Smokers or Former Smokers, Derived From the Data in Table 2 for the 3 New Studies, Stratified by Sex and Majority/US Minority Status**

	Self-Reported Smoking Status			
	Never Smokers		Former Smokers	
	Occasional	Regular	Occasional	Regular
Non-Hispanic White females	4.9% ± 1.5% (10/205)	0.6% ± 0.3% (6/977)	8.33% ± 1.9% (17/205)	0.7% ± 0.3% (7/977)
US minority females	15.6% ± 3.2% (20/128)	2.6% ± 0.6% (19/727)	5.5% ± 2.0% (7/128)	1.7% ± 0.5% (12/727)
Non-Hispanic White males	5.5% ± 3.1% (3/55)	2.4% ± 1.0% (6/251)	9.1% ± 3.9% (5/55)	3.2% ± 1.1% (8/251)
US minority males	14.5% ± 4.5% (9/62)	3.8% ± 0.9% (16/419)	16.1% ± 4.7% (10/62)	2.2% ± 0.7% (9/419)

Note. True smoking status based on cotinine level is shown in parentheses in column heads. In parentheses under each misclassification rate is the number of misclassified current smokers divided by the number of self-reported current smokers in that category. This ratio is the misclassification rate for that category. This form of misclassification rate was used in the EPA report<sup>4</sup> and is the form needed as input for the Wells-Stewart method for calculating the smoker misclassification bias.

**TABLE 4—Studies of Cotinine Levels of Nonsmokers and Smokers in the EPA Report<sup>4</sup>**

Study	Cohort Studied	Cotinine Source	Mean Cotinine Level, ng/mL, of Self-Reported Smokers	
Coultas et al. <sup>6</sup>	1043 largely Hispanic subjects aged 18+ from personal communication related to 1988 study in New Mexico	Saliva	Females Males	262 366
Cummings et al. <sup>7</sup>	663 nonsmokers aged 18-84 from personal communication related to 1986 study in Buffalo, NY	Urine	Females Males	1209 1312
Haddow et al. <sup>17</sup>	296 females examined in Portland, Me, in 1983	Serum		304
Haddow et al. <sup>18</sup>	1508 pregnant nonsmokers examined in Maine in 1984–1985 (smokers' cotinine level from 1987 study, <sup>20</sup> Table 2)	Serum		145
Lee <sup>19</sup>	808 nonsmokers aged 16–74 from 40 areas in Britain, 1985, plus cotinine levels for 176 smokers	Saliva	Females Males	264 331
Pierce et al. <sup>8</sup>	975 subjects aged 14–65+ from personal communication related to 1983 market research study in Sydney and Melbourne, Australia	Saliva	Females Males	256 <sup>a</sup> 314 <sup>a</sup>
Riboli et al. <sup>9</sup>	756 female nonsmokers aged 42–60 from personal communication to EPA related to 1986 study in China, Greece, Hong Kong, Japan, and United States by investigators who had done passive smoking studies there	Urine	Cotinine/ creatinine ratio	1200 <sup>b</sup>

<sup>a</sup>Pierce et al.<sup>8</sup> used nmol/L, which is converted to ng/mL.

<sup>b</sup>Riboli et al.<sup>9</sup> provided no information on mean cotinine levels for self-reported current smokers. A current smoker mean cotinine level of 1200 ng/mL was estimated based on Cummings et al.<sup>7</sup> (see above). The data in Dr Riboli's private communication to EPA, cited in reference 4, were in terms of ng cotinine per mg creatinine. Dr Linda Koo, who did the Hong Kong part of the Riboli study, found that her mean cotinine levels for passive smokers ranged from 10.4 to 26.6 ng/mL compared with levels of 15.6 to 39.9 ng cotinine/mg creatinine, or from 1.5 to 1.8 times smaller (private communication to EPA, December 1991). No similar information is available for the other Riboli sites. Wall et al.<sup>21</sup> found that current smokers had a mean cotinine level of 1017 ng/mL vs cotinine/creatinine of 1206 ng/mg, or 1.2 times smaller. Their passive smokers were at 9.2 ng/mL vs cotinine/creatinine of 6.3 ng/mg, or 1.5 times larger. Based on this somewhat conflicting information, the Riboli data in ng cotinine/mg creatinine were reduced by a 1.5 factor to put them on an ng/mL cotinine basis for determining the numbers of misclassified smokers.

tent with the rates for the EPA studies in Table 6, indicating that combining the rates as in Table 7 is also reasonable.

The combined misclassification rate in Table 7 for the important category of majority female regular smokers who report as never smokers is 0.8%. The combined rate for US minority female regular smokers misclassified as never smokers at 2.8% is 3.5 times the majority rate ( $P < .001$ ), and the

rate for minority occasional smokers is 2.6 times the majority rate ( $P = .001$ ). For males, the differences are similar (2.6 times,  $P = .01$ , for regular smokers and 3.9 times,  $P < .01$ , for occasional smokers). The misclassification rates for male regular smokers who report as never smokers are 1.8 times higher than the female rate for the majority category ( $P < .10$ ) and 1.3 times higher for the US minority category ( $P = .40$ ).

## Discussion

The misclassification rates for the EPA studies in Table 6 are different from those in EPA's Table B-3<sup>4</sup> because the proportion of self-reported occasional smokers to total current smokers is now 16%, based on 6 studies,<sup>6,10-12,17,19</sup> rather than EPA's estimate<sup>4</sup> of 10%, based on only 2 studies, and the 16% holds for males and females. This results in

**TABLE 5—Numbers of Misclassified Current Smokers From the Studies in the EPA Report, Stratified by Smoking Level, Sex, and Majority/US Minority Status<sup>a</sup>**

Study	Smoking Status of Misclassified Current Smokers <sup>b</sup>	Self-Reported Smoking Status, No.			Cotinine Level Ranges, ng/mL, for Misclassified Current Smokers <sup>b</sup>
		Never	Former	Current	
US non-Hispanic White females					
Cummings et al. <sup>7</sup>	Occasional	0	1	18 <sup>c</sup>	121–363
	Regular	2	0	96 <sup>c</sup>	>363
	Total subjects	225	143	114 <sup>d</sup>	
Haddow et al. <sup>17</sup>	Occasional	1 <sup>e</sup>	1 <sup>e</sup>	10	30–90
	Regular	0.5 <sup>e</sup>	0.5 <sup>e</sup>	54	>90
	Total subjects	174 <sup>f</sup>	58 <sup>f</sup>	64	
Haddow et al. <sup>18</sup>	Occasional	8 <sup>e</sup>	14 <sup>e</sup>	80 <sup>c</sup>	14–44
	Regular	1 <sup>e</sup>	1 <sup>e</sup>	423 <sup>c</sup>	>44
	Total subjects	1128 <sup>f</sup>	380 <sup>f</sup>	503 <sup>g</sup>	
Riboli et al. <sup>9</sup> (US)	Occasional	0	0	23 <sup>c</sup>	120–360
	Regular	0	0	120 <sup>c</sup>	>360
	Total subjects	224 <sup>f</sup>	81 <sup>f</sup>	143 <sup>g</sup>	
Other national majorities, females					
Lee <sup>19,h</sup> (United Kingdom)	Occasional	2 <sup>e</sup>	3 <sup>e</sup>	22	26–79
	Regular	3 <sup>e</sup>	2 <sup>e</sup>	241	>79
	Total subjects	333	125	263	
Pierce et al. <sup>8</sup> (Australia)	Occasional	2 <sup>i</sup>	2 <sup>i</sup>	27 <sup>c</sup>	26–77
	Regular	3	3	140 <sup>c</sup>	>77
	Total subjects	232	79	167	
Riboli et al. <sup>9</sup> (Shanghai, China)	Occasional	0.5 <sup>e</sup>	0.5 <sup>e</sup>	3 <sup>c</sup>	120–360
	Regular	0	0	13 <sup>c</sup>	>360
	Total subjects	94 <sup>f</sup>	5 <sup>f</sup>	16 <sup>f</sup>	
(Athens, Greece)	Occasional	0	0	2 <sup>c</sup>	120–360
	Regular	0	0	13 <sup>c</sup>	>360
	Total subjects	96 <sup>f</sup>	5 <sup>f</sup>	15 <sup>f</sup>	
(Hong Kong)	Occasional	0.5 <sup>e</sup>	0.5 <sup>e</sup>	5 <sup>c</sup>	120–360
	Regular	0	0	28 <sup>c</sup>	>360
	Total subjects	92 <sup>f</sup>	11 <sup>f</sup>	33 <sup>f</sup>	
(Sendai, Japan)	Occasional	0	0	4 <sup>c</sup>	120–360
	Regular	0.5 <sup>e</sup>	0.5 <sup>e</sup>	24 <sup>c</sup>	>360
	Total subjects	139 <sup>f</sup>	9 <sup>f</sup>	28 <sup>f</sup>	
Totals, majority females	Occasional	14	22	194	
	Regular	10	7	1152	
	Total subjects	2737	896	1346	
US minority females					
Coults et al. <sup>6</sup> (Hispanic)	Occasional	7	3	49	26–79
	Regular	5	8	135	>79
	Total subjects	387	79	184	

*(Continued)*

larger denominators for the misclassified occasional smokers and lower, somewhat more reasonable misclassification rates, whereas the regular smokers have slightly smaller denominators and slightly higher misclassification rates. This shift in percent occasional smokers has essentially no effect on the calculated smoker misclassification bias because the number of self-reported smokers of a given category multiplied by its misclassification rate is essentially unchanged.

The addition of data presented in Table 2 from Drs English and Wagenknecht (personal communications) has almost doubled the current smoker misclassification data in the very important majority female category and with data from Dr Perez-Stable (per-

sonal communication) has provided almost 6 times as much data on US minority females compared with that in the EPA report.<sup>4</sup> Majority and US minority misclassification rates are different. Therefore, in correcting passive smoking relative risks for smoker misclassification, the proportion of minorities in the particular population being studied should be considered. There appears to be little difference in the misclassification rates between males and females for occasional smokers, but there are differences (1.4% vs 0.8%) for regular smokers misclassified as never smokers, which is the most important category for calculating the bias correction.

Thus, on the basis of the current data, it would appear prudent to continue to use sex-

specific misclassification rates, at the least for majority females. The rates in Table 7 indicate that misclassification rates for occasional smokers can be quite high such that epidemiologists and prevention researchers may want to take this into account in conducting surveys on smoking behavior. Also, we were surprised that the misclassification rates for regular smokers misclassified as former smokers were on average no higher than those for regular smokers misclassified as never smokers. We had expected the former smoker rates to be higher. This similarity probably would not hold for former smokers who had recently quit.

Pregnant females who report that they are currently smoking have a lower mean

TABLE 5—Continued

Study	Smoking Status of Misclassified Current Smokers <sup>b</sup>	Self-Reported Smoking Status, No.			Cotinine Level Ranges, ng/mL, for Misclassified Current Smokers <sup>b</sup>
		Never	Former	Current	
US non-Hispanic White males Cummings et al. <sup>7,k</sup>	Occasional	1	1	19 <sup>c</sup>	131–394 >394
	Regular	1	0	97 <sup>c</sup>	
	Total subjects	116	179	116 <sup>d</sup>	
Other national majorities, males Lee <sup>19,h</sup> (United Kingdom)	Occasional	2 <sup>e</sup>	2 <sup>e</sup>	52	33–99 >99
	Regular	2 <sup>e</sup>	4 <sup>e</sup>	328	
	Total subjects	168	182	380	
Pierce et al. <sup>8,k</sup> (Australia)	Occasional	2	3	30 <sup>c</sup>	31–94 >94
	Regular	3	7	156 <sup>c</sup>	
	Total subjects	183	128	186	
Totals, majority males	Occasional	5	6	101	
	Regular	6	11	581	
	Total subjects	467	489	682	
US minority males Coults et al. <sup>6,k</sup> (Hispanic)	Occasional	6	0	14	37–110 >110
	Regular	3	8	101	
	Total subjects	171	107	115	

<sup>a</sup>As in the EPA report,<sup>4</sup> the numbers of self-reported current occasional smokers include all subjects who report themselves as current smokers but whose cotinine level is <30% of the mean cotinine level for that category as shown in Table 4.

<sup>b</sup>The cotinine level ranges for misclassified current occasional smokers are 10% to 30% of the mean levels shown in Table 4 for self-reported current smokers and >30% for misclassified regular smokers. The "smoking status" of the misclassified smokers is based on their cotinine level.

<sup>c</sup>The numbers of current occasional and regular smokers were not reported but are estimated at 16% occasional smokers and 84% regular smokers from the proportions in the other studies in Tables 2 and 5, in which the proportions were known.

<sup>d</sup>Smokers were enrolled only during the first half of the Cummings et al.<sup>7</sup> study, so the numbers of smokers in Dr Cummings' private communication were doubled to make them consistent with the other numbers.

<sup>e</sup>The split between current smokers misclassified as never smokers or former smokers was estimated for the Haddow,<sup>17,18</sup> Lee,<sup>19</sup> and Riboli<sup>9</sup> studies based on the proportions of never smokers and former smokers in the same categories from the other studies in Tables 2 and 5, in which the status was known. Specifically, from English et al.,<sup>10</sup> Wagenknecht et al.,<sup>11</sup> Cummings et al.<sup>7</sup> and Pierce et al.<sup>8</sup> for misclassified majority female occasional smokers, the proportion of never smokers was 38%; for the same category of regular smokers, the proportion was 52%; for misclassified majority male occasional smokers, the proportion was 40%; and for the same category of regular smokers, the proportion was 40%. When only 1 misclassified regular smoker was found in Haddow et al.<sup>17</sup> or Riboli et al.,<sup>9</sup> a value of 0.5 was used in both the never smoker and the former smoker columns.

<sup>f</sup>The split between never smokers and former smokers is based on national statistics.

<sup>g</sup>From the EPA report.<sup>4</sup>

<sup>h</sup>The total numbers of current smokers are calculated from the total nonsmokers and the percentages sampled as noted in the footnote to Lee's Table I. The split between never and former smokers is calculated from the ratios in Lee's Table II.

<sup>i</sup>An error in Table B-3 of the EPA report<sup>4</sup> in the numbers of misclassified female occasional smokers for Pierce et al.<sup>8</sup> is corrected in this version.

<sup>j</sup>The numbers of smokers are estimated from the numbers of never smokers and national statistics.

<sup>k</sup>The data for males in this table for Coults et al.,<sup>6</sup> Cummings et al.,<sup>7</sup> and Pierce et al.,<sup>8</sup> are from the same personal communications that were the source of data on females in the EPA report.<sup>4</sup>

cotinine level than do nonpregnant females. The misclassification rates for English et al.<sup>10</sup> and Haddow et al.<sup>18</sup> were calculated with this lower study- and category-specific smoker mean cotinine level as the basis for determining the cutpoints for misclassified occasional smokers and regular smokers. With this method, the misclassification rates for the pregnant females were not statistically significantly different from the rates for the nonpregnant females ( $P > .5$ ). Therefore, inclusion of the data from English et al.<sup>10</sup> and Haddow et al.<sup>18</sup> seemed justified.

Some investigators believe that nicotine in the diet may affect the results of this type of study.<sup>22</sup> Benowitz<sup>5</sup> has dealt with this issue and has found it to be *de minimus*. Also, an effect of different methods for ana-

lyzing cotinine is possible. Of the 10 studies, 6 used the radioimmunoassay method<sup>6,9–11,17,18</sup> and 4 used various gas or liquid chromatography methods.<sup>7,8,12,19</sup> The radioimmunoassay method is thought to be less sensitive and to give higher readings than the gas or liquid chromatography methods. However, radioimmunoassay tends to be internally consistent, and because the cotinine level of smokers from the same study is generally used to determine cutpoints and denominators (rather than fixed ng/mL levels), the misclassification rates should not be affected by analytical method.

Nine of the 10 studies reported here were done by university-based groups with epidemiologic experience. The 10th study<sup>19</sup> is statistically compatible with the other 9

( $P = .4$  for misclassified female regular smokers reporting as never smokers and ~.25 for the same category in males). Thus, the misclassification rates for the combined 10 studies appear to be of epidemiologic quality, which is a quality level required to calculate the smoker misclassification bias for epidemiologic studies of passive smoking.

## Conclusion

In conclusion, the new data on misclassification rates of smokers indicate that (1) the rates for females who report as never smokers are somewhat lower than those previously reported by EPA<sup>4</sup>; (2) investigators who plan to adjust passive smoking relative

**TABLE 6—Misclassification Rates and Standard Errors for Current Smokers Self-Reported as Either Never Smokers or Former Smokers, Derived From the Data in Table 5 for the 7 Studies in the EPA Report,<sup>4</sup> Stratified by Sex and Majority/US Minority Status**

	Self-Reported Smoking Status			
	Never Smokers		Former Smokers	
	Occasional <sup>a</sup>	Regular <sup>a</sup>	Occasional <sup>a</sup>	Regular <sup>a</sup>
Majority females	7.2% ± 1.9% (14/194)	0.9% ± 0.3% (10/1152)	11.3% ± 2.3% (22/194)	0.6% ± 0.2% (7/1152)
US minority females	14.3% ± 5.0% (7/49)	3.7% ± 1.6% (5/135)	6.1% ± 3.4% (3/49)	5.9% ± 2.0% (8/135)
Majority males	5.0% ± 2.2% (5/101)	1.0% ± 0.4% (6/581)	5.9% ± 2.4% (6/101)	1.9% ± 0.6% (11/581)
US minority males	42.9% ± 13.2% (6/14)	3.0% ± 1.7% (3/101)	0.0% (0/14)	7.9% ± 2.7% (8/101)

*Note.* In parentheses under each misclassification rate is the number of misclassified current smokers divided by the number of self-reported current smokers in that category. This ratio is the misclassification rate for that category. This form of misclassification rate was used in the EPA report<sup>4</sup> and is the form needed as input for the Wells-Stewart method for calculating the smoker misclassification bias. These misclassification rates are different from those in Table B-3 of the EPA report<sup>4</sup> because percent self-reported occasional smokers is 16% instead of 10%.

<sup>a</sup>True smoking status based on cotinine level

**TABLE 7—Misclassification Rates and Standard Errors for Current Smokers Self-Reported as Either Never Smokers or Former Smokers, Derived by Adding the Data in Table 2 for the 3 New Studies to the Data in Table 5 for the 7 EPA Studies, Stratified by Sex and Majority/US Minority Status**

	Self-Reported Smoking Status			
	Never Smokers		Former Smokers	
	Occasional <sup>a</sup>	Regular <sup>a</sup>	Occasional <sup>a</sup>	Regular <sup>a</sup>
Majority females	6.0% ± 1.2% (24/399)	0.8% ± 0.2% (16/2129)	9.8% ± 1.5% (39/399)	0.7% ± 0.2% (14/2129)
US minority females	15.3% ± 2.7% (27/177)	2.8% ± 0.6% (24/862)	5.6% ± 1.7% (10/177)	2.3% ± 0.5% (20/862)
Majority males	5.1% ± 1.8% (8/156)	1.4% ± 0.4% (12/832)	7.1% ± 2.0% (11/156)	2.3% ± 0.5% (19/832)
US minority males	19.7% ± 4.6% (15/76)	3.7% ± 0.8% (19/520)	13.2% ± 3.9% (10/76)	3.3% ± 0.8% (17/520)

*Note.* In parentheses under each misclassification rate is the number of misclassified current smokers divided by the number of self-reported current smokers in that category. This ratio is the misclassification rate for that category. This form of misclassification rate was used in the EPA report<sup>4</sup> and is the form needed as input for the Wells-Stewart method for calculating the smoker misclassification bias.

<sup>a</sup>True smoking status based on cotinine level.

risks for smoker misclassification bias should use the rates in Table 7 and should pay attention to the minority content of the cohort; (3) the use of sex-specific data is preferred; and (4) 16% should be used as the proportion of occasional smokers to total smokers, rather than the estimate of 10% used in the EPA report. In short, our work confirms the conclusion reached by EPA, namely, that the bias due to smoker misclassification is highly unlikely to be responsible for the increased risk of lung cancer observed among persons exposed to tobacco smoke. □

## References

- Lee PN, Forey BA. Misclassification of smoking habits as a source of bias in the study of environmental tobacco smoke and lung cancer. *Stat Med.* 1996;15:581-605.
- Wald NJ, Nanchahal K, Thompson SG, Cuckle HS. Does breathing other people's smoke cause lung cancer? *BMJ.* 1986;293:1217-1222.
- Lee PN. Passive smoking and lung cancer association: a result of bias? *Hum Toxicol.* 1987;6:517-524.
- Wells AJ. Appendix B: Method for correcting relative risk for smoker misclassification. In: *Respiratory Health Effects of Passive Smoking: Lung Cancer and Other Disorders: The Report of the Environmental Protection Agency,*

National Cancer Institute Monograph 4, Smoking and Tobacco Control. Washington, DC: U.S. Dept of Health and Human Services; 1993:311-335. NIH publication 93-3605.

- Benowitz NL. Cotinine as a biomarker of environmental tobacco smoke exposure. *Epidemiol Rev.* 1996;18:188-204.
- Coultas DB, Howard CA, Peake GT, Skipper BJ, Samet JA. Discrepancies between self-reported and validated cigarette smoking in a community survey of New Mexico Hispanics. *Am Rev Respir Dis.* 1988;137:810-814.
- Cummings KM, Markello SJ, Mahoney M, Bhargava AK, McElroy PD, Marshall JR. Measurement of current exposure to environmental tobacco smoke. *Arch Environ Health.* 1990;45:74-79.
- Pierce JP, Dwyer T, DiGiusto E, et al. Cotinine validation of self-reported smoking in commercially run community surveys. *J Chron Dis.* 1987;40:689-695.
- Riboli E, Preston-Martin S, Saracci R, et al. Exposure of nonsmoking women to environmental tobacco smoke: a 10-country study. *Cancer Causes Control.* 1990;1:243-252.
- English PB, Eskenazi B, Christianson RE. Black-white differences in serum cotinine levels among pregnant women and subsequent effects on infant birthweight. *Am J Public Health.* 1994;84:1439-1443.
- Wagenknecht LE, Burke GL, Perkins LL, Haley NJ, Friedman GD. Misclassification of smoking status in the CARDIA study: a comparison of self-report with serum cotinine levels. *Am J Public Health.* 1992;82:33-36.
- Perez-Stable EJ, Marin G, Marin BV, Benowitz NL. Misclassification of smoking status by self-reported cigarette consumption. *Am Rev Respir Dis.* 1992;145:53-57.
- Woodward M, Tunstall-Pedoe H, Smith WCS, Tavendale R. Smoking characteristics and inhalation biochemistry in a Scottish population. *J Clin Epidemiol.* 1991;44:1405-1410.
- Glantz SA. *Primer of Biostatistics.* New York, NY: McGraw-Hill; 1981:103.
- Perez-Stable EJ, Marin BV, Marin G, Brody DJ, Benowitz NL. Apparent underreporting of cigarette consumption among Mexican American smokers. *Am J Public Health.* 1990;80:1057-1061.
- Wagenknecht LE, Cutter GR, Haley NJ, et al. Racial differences in serum cotinine levels among smokers in the Coronary Artery Risk Development in (Young) Adults Study. *Am J Public Health.* 1990;80:1053-1056.
- Haddow JE, Palomaki GE, Knight GJ. Use of serum cotinine to assess the accuracy of self-reported non-smoking [letter]. *BMJ.* 1986;293:1306.
- Haddow JE, Knight GJ, Palomaki GE, McCarthy JE. Second-trimester serum cotinine levels in nonsmokers in relation to birth weight. *Am J Obstet Gynecol.* 1988;159:481-484.
- Lee PN. Does breathing other people's tobacco smoke cause lung cancer [letter]? *BMJ.* 1986;293:1503-1504.
- Haddow JE, Knight GJ, Palomaki GE, Kloza EM, Wald NJ. Cigarette consumption and serum cotinine in relation to birthweight. *Br J Obstet Gynaecol.* 1987;94:678-681.
- Wall MA, Johnson J, Jacob P, Benowitz NL. Cotinine in the serum, saliva, and urine of non-smokers, passive smokers and active smokers. *Am J Public Health.* 1988;78:699-701.
- Davis RA, Stiles MF, deBethizy JD, et al. Dietary nicotine: a source of urinary cotinine. *Food Chem Toxicol.* 1991;29:821-827.