

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications, Department of Psychology

Psychology, Department of

---

2014

## Consistency and Recanting of Ever-Smoking Status Reported by Self and Proxy Respondents One Year Apart

Julia N. Soulakova

University of Nebraska-Lincoln, Julia.Soulakova@unl.edu

Lisa J. Crockett

University of Nebraska-Lincoln, lcrockett1@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/psychfacpub>



Part of the [Psychology Commons](#)

---

Soulakova, Julia N. and Crockett, Lisa J., "Consistency and Recanting of Ever-Smoking Status Reported by Self and Proxy Respondents One Year Apart" (2014). *Faculty Publications, Department of Psychology*. 758.

<https://digitalcommons.unl.edu/psychfacpub/758>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in final edited form as:

*J Addict Behav Ther Rehabil.* ; 3(1): 1000113–. doi:10.4172/2324-9005.1000114.

# Consistency and Recanting of Ever-Smoking Status Reported by Self and Proxy Respondents One Year Apart

Julia N Soulakova<sup>1,\*</sup> and Lisa J Crockett<sup>2</sup>

<sup>1</sup>Department of Statistics, University of Nebraska-Lincoln, USA

<sup>2</sup>Department of Psychology, University of Nebraska-Lincoln, USA

## Abstract

The study examined consistency of reports concerning current and prior smoking behaviors. Data came from the 2002–2003 Tobacco Use Supplement to the Current Population Survey, where the current smoking behaviors and smoking history were reported by self- and proxy-respondents on two occasions, one year apart. The ever-smoking status is reported consistently, overall: Kappa coefficient is 0.78 with the corresponding 95% confidence interval given by (0.77, 0.78). One specific type of inconsistency of prior reports was assessed for respondents who were identified as never smokers at the latter assessment and former or current smokers at the earlier assessment. Based on the survey logistic regression that controls for multiple respondent characteristics and survey administration method, the estimated prevalence of such inconsistent self-reports is 9.0%, and prevalence of inconsistent proxy-reports is 5.4%. In addition, prevalence of recanting, i.e., future reporting never smoking for respondents who previously claimed to be a former or a current smoker was assessed. The recanting was shown to be most prevalent with respect to proxy-reports and former smokers: overall prevalence of recanting was estimated to be in the range 13% – 19% for current smokers, and 27% – 46% for former smokers. Our findings indicate that while, unexpectedly, proxy-respondents are more likely to report the ever-smoking status consistently than do self-respondents, the proxies are also more likely to incorrectly report never smoking in the future for smokers especially regarding adolescents and young adults. Therefore, the observed higher level of consistency for proxy-respondents may be due to proxies' incorrect knowledge which leads to consistent yet ambiguous responses.

## Keywords

Data agreement; Complex survey; Reliability; Respondent type; Survey logistic regression

## Introduction

### Background

Ever-smoking status, i.e., smoking at least 100 cigarettes in the life-time, is the core information in the tobacco use research, and is assessed in many national surveys, e.g. the National Health Interview Survey (NHIS) [1], and the National Longitudinal Surveys (NLS) [2]. The ever-smoking status provides the means for defining one's current smoking status as never smoker, every day smoker, someday smoker and former smoker, as is presented in the Appendix. Self-reported ever-smoking status, as well as several other smoking measures assessed in the surveys has been shown to be reliable [3–6].

As was previously pointed out [6,7], investigating consistency of ever-smoking reports is not as straightforward as in the case of some other smoking-related information which is expected to remain constant with time, e.g., age of smoking initiation. If the smoking-related information is not anticipated to change with time, strict consistency can be assessed using the prevalence of exact matches in the reports on two (or more) occasions. In contrast, because some subjects may start smoking between the surveys, their ever-smoking status may change with time, and for these subjects non-matching responses should not be treated as inconsistent ones. Thus, it is important to carefully identify truly inconsistent responses and refine the corresponding consistency measures.

Such a careful differentiation between the consistent and inconsistent reports was used in a study [7], where recanting (denial of previous smoking reports) was defined and the patterns of recanting was examined across different smoking stages (never, puffer, light, occasional, regular, ex-smoker). The study used the adolescent and young adult reports to the 1995 and 2001–2002 National Longitudinal Study of Adolescent Health surveys. The study investigated potential influence of age, gender, race/ethnicity, poverty status, and other factors on the prevalence of recanting. It was shown that the prevalence of recanting is higher in respondents who indicated less frequent/intense smoking behaviors in 1995. In addition, it was shown that male reports may be less reliable than female reports, and minority youth respondents may be more likely to recant their smoking behaviors than do white youth respondents. Moreover, the authors demonstrated that the relationship between the prevalence of recanting and other factors may be quite complex. Test-retest reliability of smoking behaviors in adolescents and young adults, and influence of demographic and other factors on the reliability level was also investigated in several other studies [8–16].

Since many national surveys allow interviewing proxy-respondents (e.g., partners, parents, friends) if the intended subjects cannot be surveyed, confirming validity and reliability of proxy-reports is also essential. Indeed, while inclusion of proxy-respondents helps reduce survey costs and increase the response rates, inclusion of proxy-reports may also lead to under-reporting the smoking behaviors [17,18]. In addition, the data quality may depend on the interview method (phone, in-person) [18–20], race/ethnicity [21], as well as the joint effects of respondent and survey characteristics [6,7,21].

## Data and study objectives

In this paper we address consistency of self- and proxy-reported ever-smoking information in the 2002–2003 Tobacco Use Supplement to the Current Population Survey (TUS-CPS) that had two survey waves, in 2002 and 2003. The TUS-CPS is sponsored by the U.S. Bureau of Labor Statistics and fielded by the U.S. Census Bureau [22]. The surveys were administered to self- or proxy-respondents using a phone or in-person interviews: the “phone” interview method refers to the interviews conducted over a phone both times, “in-person” interview method refers to the interviews conducted once in-person and once over a phone; the “self” respondent type refers to the instances when self-respondents were interviewed both times, the “proxy” respondent type refers to the instances when proxy-respondents were interviewed both times, and “self-proxy” respondent type refers to the instances when self-respondents were interviewed once and proxy-respondents were interviewed once.

Our first goal is to assess specific type of consistency in the ever-smoking reports. For this purpose we limit our consideration to respondents who claimed (themselves or by proxies) to be never smokers in 2003. Then respondents who claimed (themselves or by proxies) to be never smokers in 2002 are said to provide consistent responses, while the respondents who claimed (themselves or by proxies) to be former or current smokers in 2002 are said to provide inconsistent responses. For simplicity, this measure is called the ever-smoking status (consistency). We identify the relationship between the prevalence (*logit*) of consistent reporting of ever-smoking and key intended subjects’ respondent and survey characteristics. Next based on the identified relationship we test whether the prevalence of consistent responses depends on (1) the respondent type (self, proxy, self-proxy), (2) the joint effect of the respondent type and the intended subjects’ age (15–24, 25–44, 45–64, 65+), (3) the joint effect of the respondent type and intended subjects’ gender (male, female), and (4) the joint effect of the respondent type and interview method (phone, in-person, mixed). Finally, if any of the joint effects are identified to be significant, then within each particular subpopulation we assess whether self-reporting leads to different prevalence of consistent responses when compared to proxy reporting and self-proxy reporting.

Our second goal is to assess another specific type of disagreement, here called recanting. As is mentioned in the Section 1, self-reports and in-person interviews are expected to result in accurate and valid current smoking information, this is why we define recanters as subjects who self-reported being a current or former smoker in 2002 via in-person interviews but who later (2003) claimed never to have smoked (via self-report or proxy report). First, the overall relationship between the prevalence (*logit*) of future reporting never smoking and key characteristics are identified based on self-reports done in-person in 2002. Then the relationship will be used to assess whether the prevalence of future reporting never smoking for recanters depends on (1) the joint effect of the respondent type (self, proxy), intended subjects’ smoking status (current smoker, former smoker) and age; (2) joint effect of the respondent type, intended subjects’ smoking status and gender; and (3) joint effect of the respondent type, intended subjects’ smoking status and the interview method (phone, in-person); where the respondent type and interview method refer to the 2003 wave, and current smoking status refers to the 2002 smoking status. Finally, we test whether the self-

reporting results in different prevalence of future reporting never smoking than does proxy-reporting within each specific subpopulation of recanters (provided there is a significant corresponding joint effect).

## Materials and Methods

We used the Balanced Repeated Replications approach with replicate weights [23] to properly adjust for the complex design of the TUS-CPS. The BRR was used to produce all data summary statistics, Kappa coefficients, as well as logistic regression models. We used SAS/STAT® 9.2 software [24] and SUDAAN® 11 software [25] to perform the statistical analyses.

There are 22,255 subjects who reported their smoking status in 2002 and 2003 and have complete information with respect to all characteristics of interest. Tables 1 and 2 illustrate the weighted summary statistics for the samples corresponding to ever-smoking and future reporting never-smoking measures, respectively.

The strict data agreement was measured using Kappa coefficient. Kappa is an over-conservative measure because it incorrectly allocates respondents who started smoking between the surveys into “inconsistent respondent” group. Since we believe that Kappa should be used for explorative purposes only, we do not perform any specific testing across the respondent types in terms of Kappa measures; the reported intervals should not be mistreated as the 95% simultaneous confidence bands.

The *logit* of consistent ever-smoking status responses is modeled as a function of the baseline demographic characteristics of the intended subjects, i.e., age, gender, race/ethnicity, metropolitan status and region, as well as the interview method (phone, in-person, mixed), and respondent type (self, proxy, self-proxy). The initial model contained all two-way interactions (in addition to the main effects). Then the back-ward elimination approach was used to exclude all insignificant (at 5% level) two-way interactions provided the interaction did not correspond to the relationship of interest; interactions corresponding to the relationships of interest were kept in the model despite their insignificance.

The *logit* of future reporting never smoking was modeled as a function of the baseline demographic characteristics (mentioned above), current smoking status, and 2003 respondent type and 2003 interview method, as well as all two-way and specific three-way interactions. Since our primary interest is to identify the factors associated with recanting, the three-way interactions among (a) age, smoking status and respondent type, (b) gender, smoking status and respondent type, and (c) interview method, smoking status and respondent type were included, where smoking status corresponds to the 2002 current smoking status which is recorded as never smoker, smoker (someday or everyday) and former smoker; and the respondent type and interview method refer to the 2003 assessment. This model was then used to assess objectives corresponding to the recanting.

Based on the final models, the specific objectives stated in Section 2 with respect to ever-smoking status and recanting were investigated individually, each at an overall 5% significance level. In the case of a significant joint effect of interest, the corresponding

planned comparisons were performed using the Bonferroni adjustments for multiplicity. All hypotheses were tested against the corresponding two-sided alternatives stating that there are differences across the subpopulations, so all reported p-values correspond to the two-sided tests.

## Results

Out of 8,967 respondents who were ever-smokers in 2002, 82.3% were ever-smokers in 2003, while 17.7% were never smokers in 2003. In addition, out of 13,288 who were never-smokers in 2002, 6.1% were ever-smokers and 93.9% were never-smokers in 2003. Also, we observed that out of 1,358 respondents who self-reported to be a current smoker in 2002, the percent of recanting is 13.5%, while out of 1,640 respondents who self-reported to be a former smoker in 2002, the percent of recanting is higher, that is 25.1%,

As is expected, the overall percentage of consistent responses is relatively high for the ever-smoking status, i.e., 89.9% (Table 1). The Kappa statistics also confirm relatively high level of data consistency: the Kappa coefficient is 0.7736 (SE=0.0028; 95% CI=(0.7681, 0.7790)). The Kappa coefficients corresponding to self, proxy and self-proxy subpopulations are given, respectively, by 0.7909 (SE=0.0032; CI=0.7844, 0.7973), 0.7555 (SE=0.0106; CI=0.7345, 0.7765), and 0.7152 (0.0066, CI=0.7021, 0.7283).

### Ever-smoking status

The final models corresponding to the ever-smoking status measure contains a large number of two-way interactions (in addition to all main effects), the model is significant at 5% level (Chi-Square=3,915,938, df=83,  $p<0.0001$ ).

**Objective 1**—There are multiple significant interactions that involve the respondent type (i.e., with age, race/ethnicity, metropolitan status, and region), and the main effect is also significant ( $p<0.0001$ ). Therefore, we conclude that the prevalence of consistent responses depends on the respondent type. This model-based conclusion is also supported by the Rao-Scott test for independence (Wald  $F(2, 79)=33.1$ ,  $p<0.0001$ ): self, proxy and self-proxy respondents do differ significantly in terms of the prevalence of consistent responses. As is depicted in Table 3 the model results in somewhat unexpected patterns: the predicted prevalence of consistent responses is 91.0% for self-respondents, 94.6% for proxy-respondents, and 91.7% for self-proxy respondents. Non-modeled data result in the similar pattern: the observed weighted proportions are 89.5% for self-respondents, 92.9% for proxy-respondents, and 89.8% for self-proxy respondents. The differences are significant between the prevalence of self- and proxy-reports (Chi-square=26.0, df=1,  $p<0.0001$ ), proxy and self-proxy reports (Chi-square=14.5, df=1,  $p=0.0001$ ), but there is no significant difference between self and self-proxy reports (Chi-square= 1.4, df=1,  $p=0.2360$ ).

**Objective 2**—There is a significant two-way interaction between the respondent type and age group (Chi-square=49.2, df=6,  $p<0.0001$ ). Planned comparisons between the respondent types (each at 0.625% significance level) result in the following conclusions: the 15–24 year old self-respondents are less likely to provide consistent responses than are the proxy and

self-proxy respondents, in addition the 25–64 year old self-respondents are less likely to provide consistent responses than are the proxy respondents.

**Objectives 3 and 4**—The interactions between the respondent type and gender ( $p=0.7401$ ) and the respondent type and interview method ( $p$ -value is 0.1472) are not significant, after controlling for the other covariates. Therefore, the differences in the proportions associated with the respondent type are similar for females and males, and across the three interview methods.

### Future reports of never smoking; recanting

The final model (Chi-Square= 41,683,  $df=93$ ,  $p<0.0001$ ) contains three significant three-way interactions: the interaction between the respondent type, current smoking status and age (Chi-Square=26.67,  $df=6$ ,  $p=0.0002$ ), respondent type, current smoking status and gender (Chi-Square=9.32,  $df=2$ ,  $p=0.0095$ ), and respondent type, current smoking status, and interview method (Chi-square=6.46,  $df=2$ ,  $p=0.0396$ ). Table 4 presents the predicted proportions of future reporting never-smoking with respect to each subpopulation of interest.

The prevalence of future reporting never smoking depends on the joint effect of the respondent type and current smoking status. The non-modeled weighted sample proportions of reporting never-smoking for 2002 current smokers are 0.13 ( $SE=0.01$ ) and 0.16 ( $SE=0.02$ ) for self- and proxy-reports, respectively; and for 2002 former smokers the proportions are 0.24 ( $SE=0.01$ ) and 0.33 ( $SE=0.02$ ) for self- and proxy-reports, respectively. The model results in the similar patterns, depicted in Table 4: while self-respondents are less likely to recant than are proxy-respondents, the odds ratio is significant only for 2002 former smokers. Also note that the prevalence of recanting for 2002 former smokers is higher than it is for 2002 current smokers, the result is consistent across the two respondent types.

**Objective 1**—Self-respondents (who indicated current or former smoking in 2002) of any age are less likely to recant than do proxy respondents. However, the differences in prevalence of recanting are significant only for 45–64 subjects who were current smokers in 2002 and 25–44 subjects who were former smokers in 2002.

**Objective 2**—Significant differences in prevalence of recanting correspond to females who were current smokers in 2002 and males and females who were former smokers in 2002: within each subpopulation the self-respondents are less likely to recant than are proxy-respondents.

**Objective 3**—Significant differences in prevalence of recanting correspond to respondents who were current smokers in 2002 and whose smoking status was reported via a phone interview in 2003 and respondents who were former smokers in 2002 and whose smoking status was reported via a phone or in-person interview in 2003: within each interview method and prior smoking status subpopulation the self-respondents are less likely to recant than are proxy-respondents.



## Discussion

We observed that the ever-smoking is reported, overall, consistently. However, the prevalence of consistent responses varies across the respondent types and depends on the age of intended subjects. There is an unexpected finding that proxy-respondents are more likely, overall, to provide consistent responses than do self-respondents. However, our additional analyses concerning the prevalence of recanting indicated that proxy-respondents are actually more likely to incorrectly report subjects' never smoking in the future than do the self-respondents, and the differences also depend on other characteristics. The consistent finding was that the largest prevalence of recanting is observed with respect proxy-reports rather than self-reports, and concerned individuals whose prior smoking status was a former smoker rather than current smoker. This agrees with related findings that denials are more common in occasional smokers rather than everyday smokers [7,26].

Our findings with respect to ever-smoking status and recanting may suggest that although proxies are more likely to provide consistent responses, they are also more likely to provide inaccurate information than do self-respondents, thus validity of proxy-reports is questionable. Furthermore, the extent of the discrepancies between the self- and proxy-reports may depend on other factors.

The study has several limitations. The planned comparisons were model-based. Since the models may be, potentially, improved to better fit the data, the corresponding new estimates may lead to different conclusions. However, since the presented non-model and model based estimates agree, we do not anticipate any drastic changes in the conclusions. Also, we examined only the prevalence of consistent responses provided one-year apart. The larger time periods between the surveys may result in smaller prevalence of consistent responses, as was previously observed in other studies [3,7].

## Acknowledgments

Research of Julia N. Soulakova reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R03CA165831. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors wish to thank Anne Hartman and Todd Gibson (National Cancer Institute) for providing the data set.

## References

1. Centers for Disease Control and Prevention. National Health Interview Survey. 2013
2. U.S. Bureau of Labor Statistics. National Longitudinal Surveys. 2013
3. Kenkel D, Lillard DR, Mathios A. Smoke or fog? The usefulness of retrospectively reported information about smoking. *Addiction*. 2003; 98:1307–1313. [PubMed: 12930218]
4. Brigham J, Lessov-Schlaggar CN, Javitz HS, McElroy M, Krasnow R, et al. Reliability of adult retrospective recall of lifetime tobacco use. *Nicotine Tob Res*. 2008; 10:287–299. [PubMed: 18236293]
5. Colby SM, Clark MA, Rogers ML, Ramsey S, Graham AL, et al. Development and reliability of the lifetime interview on smoking trajectories. *Nicotine Tob Res*. 2012; 14:290–298. [PubMed: 21994340]
6. Soulakova JN, Hartman AM, Liu B, Willis GB, Augustine S. Reliability of adult self-reported smoking history: data from the tobacco use supplement to the current population survey 2002–2003 cohort. *Nicotine Tob Res*. 2012; 14:952–960. [PubMed: 22318688]



7. Stanton CA, Papandonatos G, Lloyd-Richardson EE, Niaura R. Consistency of self-reported smoking over a 6-year interval from adolescence to young adulthood. *Addiction*. 2007; 102:1831–1839. [PubMed: 17784897]
8. Raghupathy S, Hahn-Smith S. Reliability of the high school Youth Risk Behavior Survey when administered online. *Int Q Community Health Educ*. 2011; 32:135–148. [PubMed: 23000460]
9. Gillum RF. Religiosity and the validity of self-reported smoking: the Third National Health and Nutrition Examination Survey. *Review of Religious Research*. 2005; 47:190–196.
10. Huerta M, Chodick G, Balicer RD, Davidovitch N, Grotto I. Reliability of self-reported smoking history and age at initial tobacco use. *Prev Med*. 2005; 41:646–650. [PubMed: 15917064]
11. Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EC, et al. Reliability of the 1999 youth risk behavior survey questionnaire. *J Adolesc Health*. 2002; 31:336–342. [PubMed: 12359379]
12. Johnson TP, Mott JA. The reliability of self-reported age of onset of tobacco, alcohol and illicit drug use. *Addiction*. 2001; 96:1187–1198. [PubMed: 11487424]
13. Nelson DE, Holtzman D, Bolen J, Stanwyck CA, Mack KA. Reliability and validity of measures from the Behavioral Risk Factor Surveillance System (BRFSS). *Soz Praventivmed*. 2001; 46(Suppl 1):S3–S42. [PubMed: 11851091]
14. Henriksen L, Jackson C. Reliability of children's self-reported cigarette smoking. *Addict Behav*. 1999; 24:271–277. [PubMed: 10336108]
15. Stanton WR, McClelland M, Elwood C, Ferry D, Silva PA. Prevalence, reliability and bias of adolescents' reports of smoking and quitting. *Addiction*. 1996; 91:1705–1714. [PubMed: 8972928]
16. Brener ND, Collins JL, Kann L, Warren CW, Williams BI. Reliability of the Youth Risk Behavior Survey Questionnaire. *Am J Epidemiol*. 1995; 141:575–580. [PubMed: 7900725]
17. Harakeh Z, Engels RC, Vries Hd, Scholte RH. Correspondence between proxy and self-reports on smoking in a full family study. *Drug Alcohol Depend*. 2006; 84:40–47. [PubMed: 16386380]
18. Soulakova J, Davis WW, Hartman A, Gibson J. The Impact of Survey and Response Modes on Current Smoking Prevalence Estimates Using TUS-CPS: 1992–2003. *Surv Res Methods*. 2009; 3:123–137. [PubMed: 21841957]
19. Simile, C.; Stussman, B.; Dahlhamer, J. Exploring the impact of mode on key health estimates in the National Health Interview Survey; Proceedings of Statistics Canada Symposium 2006: Methodological Issues in Measuring Population Health; 2006.
20. Gundersen, D.; Delnevo, C.; Momperousse, D. Mode effects on cigarette smoking estimates: comparing CAPI and CATI responders in the 2001/02 Current Population Survey. 135th Annual Meeting of the American Public Health Association; Washington, DC. 2007.
21. Navarro AM. Smoking status by proxy and self report: rate of agreement in different ethnic groups. *Tob Control*. 1999; 8:182–185. [PubMed: 10478403]
22. U.S. Department of Health and Human Services. National Cancer Institute and Centers for Disease Control and Prevention Co-sponsored Tobacco Use Supplement to the Current Population Survey. 2013
23. Davis W, Hartman A, Gibson J. Weighting the overlap sample obtained from two Tobacco Use Supplements to the Current Population Survey. 2007
24. SAS Institute Inc. SAS/STAT® 9.2 User's Guide. Cary, NC: SAS Institute Inc.; 2008.
25. Research Triangle Institute. SUDAAN Language Manual. Vol. Volumes 1 and 2. Research Triangle Park, NC: Research Triangle Institute; 2012. Release 11
26. Cowling DW, Johnson TP, Holbrook BC, Warnecke RB, Tang H. Improving the self reporting of tobacco use: results of a factorial experiment. *Tob Control*. 2003; 12:178–183. [PubMed: 12773728]

**Table 1**

Sample corresponding to ever-smoking consistency measure: sample count and percentage (in parentheses) corresponding to the population count (provided in the Overall row).

	Self Respondents	Proxy Respondents	Self-proxy Respondents	Overall
<b>Ever-Smoking Status</b>				
Consistent Responses	8444 (89.5%)	1248 (92.9%)	2781 (89.8%)	12473 (89.9%)
Inconsistent Responses	1022 (10.5%)	108 (7.1%)	358 (10.2%)	1488 (10.1%)
<b>Age</b>				
15–24	991 (16.0%)	517 (46.0%)	799 (34.1%)	2307 (23.8%)
25–44	3656 (40.8%)	436 (31.0%)	1147 (35.4%)	5239 (38.4%)
45–64	3020 (27.1%)	277 (15.4%)	928 (23.5%)	4225 (24.9%)
65+	1799 (16.1%)	126 (7.6%)	265 (6.9%)	2190 (12.9%)
<b>Gender</b>				
Male	3487 (38.9%)	777 (57.2%)	1634 (52.2%)	5898 (44.2%)
Female	5979 (61.1%)	579 (42.8%)	1505 (47.8%)	8063 (55.8%)
<b>Race/Ethnicity</b>				
Non-Hispanic White	7409 (67.9%)	973 (63.6%)	2297 (63.7%)	10679 (66.4%)
Other	2057 (32.1%)	383 (36.4%)	842 (36.3%)	3282 (33.6%)
<b>Metropolitan Status</b>				
Metropolitan	6861 (80.1%)	1067 (85.2%)	2403 (83.2%)	10331 (81.4%)
Non-Metropolitan	2605 (19.9%)	289 (14.8%)	736 (16.8%)	3630 (18.6%)
<b>Region</b>				
Northeast	1869 (17.2%)	367 (24.7%)	684 (19.9%)	2920 (18.7%)
Midwest	2540 (22.4%)	312 (19.4%)	754 (21.0%)	3606 (21.7%)
South	2788 (36.8%)	391 (34.0%)	953 (36.0%)	4132 (36.3%)
West	2269 (23.5%)	286 (22.0%)	748 (23.1%)	3303 (23.3%)
<b>Interview Method</b>				
Phone both times	5213 (53.2%)	776 (56.0%)	1605 (48.8%)	7594 (52.5%)
In-person both times	2267 (25.1%)	274 (20.8%)	719 (24.8%)	3260 (24.6%)
Mixed	1986 (21.6%)	306 (23.1%)	815 (26.4%)	3107 (23.0%)
Overall	9466 (64.2%) 93,000,797	1356 (11.4%) 16,548,939	3139 (24.3%) 35,204,079	13961 (100%) 144,753,815

**Table 2**

Sample corresponding to future reporting never smoking measure: sample count and percentage (in parentheses) corresponding to the population count (provided in the Overall row).

	<b>2003 Respondent Type</b>		
	<b>Self Respondents</b>	<b>Proxy Respondents</b>	<b>Overall</b>
<b>2002 Smoking Status</b>			
Never Smoker	3365 (58.7%)	644 (66.9%)	4009 (60.2%)
Current Smoker	1179 (19.0%)	179 (17.3%)	1358 (18.7%)
Former Smoker	1444 (22.1%)	196 (15.8%)	1640 (21.1%)
<b>2003 Never-Smoking Status</b>			
Never Smoker	3656 (63.5%)	692 (71.4%)	4348 (60.2%)
Current or Former Smoker	2332 (36.5%)	327 (28.6%)	2659 (18.7%)
<b>Age</b>			
15–24	518 (13.4%)	216 (29.8%)	734 (16.0%)
25–44	2161 (38.4%)	366 (36.3%)	2527 (38.1%)
45–64	2100 (30.4%)	331 (25.5%)	2431 (29.6%)
65+	1209 (17.8%)	106 (8.4%)	1315 (16.3%)
<b>Gender</b>			
Male	2521 (44.3%)	569 (56.6%)	3090 (46.3%)
Female	3467 (55.7%)	450 (43.4%)	3917 (53.7%)
<b>Race/Ethnicity</b>			
Non-Hispanic White	4570 (68.7%)	756 (65.3%)	5426 (68.2%)
Other	1318 (31.3%)	263 (34.7%)	1581 (31.8%)
<b>Metropolitan Status</b>			
Metropolitan	4382 (80.0%)	772 (82.0%)	5154 (80.4%)
Non-Metropolitan	1606 (20.0%)	247 (18.0%)	1853 (19.6%)
<b>Region</b>			
Northeast	1247 (18.4%)	233 (21.7%)	1480 (18.9%)
Midwest	1510 (21.9%)	252 (21.9%)	1762 (21.9%)
South	1819 (37.0%)	289 (33.4%)	2108 (36.4%)
West	1412 (22.7%)	245 (22.9%)	1657 (22.7%)
<b>2003 Interview Method</b>			
Phone	2274 (37.5%)	472 (44.5%)	2726 (38.6%)
In-person	3714 (62.5%)	547 (55.5%)	4261 (61.4%)
Overall	5988 (84.0%) 58,961,951	1019 (16.0%) 11,234,338	7,007 (100%) 70,196,289

**Table 3**

Model-based predicted proportions of consistent responses and the corresponding standard errors, overall odds ratios of consistent responses with the corresponding standard errors (top entry) and p-values (bottom entry).

	Prevalence of consistent responses across 2002 Respondent Type		Odds ratios of consistent responses	
	Self	Proxy	Self-proxy	Self versus Proxy Self versus Self-proxy
<b>Overall</b>	0.91 ± 0.00	0.95 ± 0.01	0.92 ± 0.01	<b>0.57 ± 0.06*</b> 0.91 ± 0.07 0.2360
Age (adjusted individual significance level is 0.625%)				
15–24	0.96 ± 0.01	0.99 ± 0.00	0.98 ± 0.00	<b>0.26 ± 0.05*</b> 0.56 ± 0.09 0.0002
25–44	0.91 ± 0.00	0.93 ± 0.01	0.90 ± 0.01	<b>0.71 ± 0.08</b> <b>0.0031</b> 1.04 ± 0.08 0.6411
45–64	0.86 ± 0.01	0.91 ± 0.01	0.87 ± 0.01	<b>0.57 ± 0.08*</b> 0.92 ± 0.08 0.3361
65+	0.88 ± 0.01	0.88 ± 0.02	0.85 ± 0.02	1.03 ± 0.18 0.8730 1.30 ± 0.16 0.0311
<b>Gender</b>				
Male	0.88 ± 0.01	0.93 ± 0.01	0.89 ± 0.01	0.56 ± 0.07* 0.89 ± 0.08 0.1697
Female	0.93 ± 0.00	0.96 ± 0.01	0.94 ± 0.00	0.58 ± 0.08* 0.94 ± 0.08 0.4498
<b>Interview Method</b>				
Phone	0.92 ± 0.00	0.95 ± 0.01	0.92 ± 0.01	0.67 ± 0.09 0.0018 1.04 ± 0.09 0.6753
In-Person	0.89 ± 0.01	0.94 ± 0.01	0.91 ± 0.01	0.50 ± 0.09* 0.80 ± 0.08 0.0236
Mixed	0.91 ± 0.01	0.95 ± 0.01	0.92 ± 0.01	0.56 ± 0.08* 0.92 ± 0.10 0.4234

**Note:** Significant (after multiplicity adjustments) results corresponding to tested hypotheses are highlighted in bold;

\* indicates the p-value less than or equal to 0.0001.

**Table 4**

Future reporting never smoking: model-based predicted proportions (with standard errors), odds ratios with standard errors and corresponding p-values.

	Current Smoker in 2002			Former Smoker in 2002		
	2003 Respondent Type		Self versus Proxy Odds Ratio	2003 Respondent Type		Self versus Proxy Odds Ratio
	Self	Proxy		Self	Proxy	
<b>Overall</b>	0.13 ± 0.01	0.19 ± 0.04	0.60 ± 0.15 0.0392	0.27 ± 0.01	0.46 ± 0.07	<b>0.42 ± 0.11</b> <b>0.0011</b>
<b>Age (adjusted individual significant level is 0.625%)</b>						
15–24	0.11 ± 0.02	0.11 ± 0.05	0.99 ± 0.51 0.9781	0.26 ± 0.04	0.73 ± 0.17	0.13 ± 0.12 0.0208
25–44	0.14 ± 0.01	0.16 ± 0.03	0.87 ± 0.24 0.6044	0.31 ± 0.02	0.49 ± 0.05	<b>0.47 ± 0.10</b> <b>0.0003</b>
45–64	0.09 ± 0.01	0.29 ± 0.05	<b>0.25 ± 0.05*</b>	0.26 ± 0.02	0.33 ± 0.04	0.71 ± 0.13 0.0706
65+	0.17 ± 0.03	0.25 ± 0.11	0.61 ± 0.32 0.3484	0.24 ± 0.02	0.31 ± 0.06	0.71 ± 0.20 0.2231
<b>Gender (adjusted individual significant level is 1.25%)</b>						
Male	0.13 ± 0.01	0.12 ± 0.03	1.09 ± 0.28 0.7200	0.26 ± 0.01	0.46 ± 0.07	<b>0.42 ± 0.12</b> <b>0.0026</b>
Female	0.12 ± 0.01	0.30 ± 0.07	<b>0.33 ± 0.11</b> <b>0.0006</b>	0.28 ± 0.02	0.47 ± 0.07	<b>0.42 ± 0.12</b> <b>0.0028</b>
<b>2003 Interview Method (adjusted individual significant level is 1.25%)</b>						
Phone	0.11 ± 0.02	0.27 ± 0.06	<b>0.35 ± 0.10</b> <b>0.0002</b>	0.26 ± 0.02	0.47 ± 0.07	<b>0.40 ± 0.11</b> <b>0.0005</b>
In-Person	0.14 ± 0.01	0.13 ± 0.04	1.05 ± 0.33 0.8890	0.27 ± 0.01	0.45 ± 0.08	<b>0.45 ± 0.14</b> <b>0.0098</b>

**Note:** Significant (after multiplicity adjustments) results corresponding to tested hypotheses are highlighted in bold;

\* indicates the p-value less than or equal to 0.0001.