Assessment of validity of self-reported smoking status

by Suzy L. Wong, Margot Shields, Scott Leatherdale, Eric Malaison and David Hammond

Abstract

Background
Cigarette smoking is associated with adverse health effects, including cancer, respiratory illness, heart disease and stroke. National data on smoking prevalence often rely on self-reports. This study assesses the validity of self-reported cigarette smoking status among Canadians.

Data and methods
Data are from the 2007 to 2009 Canadian Health Measures Survey, a nationally representative cross-sectional survey of 4,530 Canadians aged 12 to 79. The survey included self-reported smoking status and a measure of urinary cotinine, a biomarker of exposure to tobacco smoke. The prevalence of cigarette smoking was calculated based on self-reports and also on urinary cotinine concentrations.

Results
Compared with estimates based on urinary cotinine concentration, smoking prevalence based on self-report was 0.3 percentage points lower. Sensitivity estimates (the percentage of respondents who reported being smokers among those classified as smokers based on cotinine concentrations) were similar for males and females (more than 90%). Although sensitivity tended to be lower for respondents aged 12 to 19 than for those aged 20 to 79, the difference did not attain statistical significance.

Interpretation
Accurate estimates of the prevalence of cigarette smoking among Canadians can be derived from self-reported smoking status data.

Keywords
Biological specimens, cotinine, data collection, direct measures, health surveys, reproducibility of results, urine specimen collection

Authors
Suzy L. Wong (613-951-4774; suzy.wong@statcan.gc.ca) and Margot Shields (613-951-4177; margot.shields@statcan.gc.ca) are with the Health Analysis Division at Statistics Canada, Ottawa, Ontario, K1A 0T6. Scott Leatherdale and David Hammond are with the University of Waterloo, Waterloo, Ontario. Eric Malaison is with Health Canada.

The health risks associated with cigarette smoking are well-documented and widely recognized—cancer, respiratory illness, heart disease, and stroke.¹⁻³ In Canada, smoking contributes to more than 37,000 deaths a year.⁴ Tobacco-related health care expenditures amount to billions of dollars annually, with additional indirect costs such as lost productivity, longer-term disability and premature death.⁴

Self-reported data are typically used to monitor trends in cigarette smoking.⁵⁻⁷ However, estimates based on self-report, particularly of socially undesirable behaviours, are subject to reporting biases.⁸ The widespread implementation of legislation prohibiting smoking in workplaces and public areas⁹ and prominent health warnings on cigarette packages may reinforce the perception of smoking as socially undesirable, and thereby increased the tendency to underreport over time.

To validate self-reported smoking status, the urinary concentration of cotinine, a widely accepted objective measure of exposure to tobacco smoke,¹⁰ has been used. Cotinine is the major metabolite of nicotine, with a half-life of about 16 to 20 hours.¹¹ Because of its high sensitivity and specificity, cotinine is considered to be an accurate quantitative measure of recent exposure to tobacco smoke.¹² Compared with estimates based on cotinine concentration, smoking prevalence based on self-report, is generally lower,¹³ although the extent of the difference varies by country.¹⁴

The validity of self-reported cigarette smoking data have yet to be determined for Canada. Thus, this study compares estimates of the prevalence of cigarette smoking based on self-report with estimates based on urinary cotinine concentrations. The data are from the 2007 to 2009 Canadian Health Measures Survey, which included self-reported smoking status and the first nationally representative measures of urinary cotinine.
Methods

Data source
The Canadian Health Measures Survey (CHMS) is a nationally representative survey of the household population. Data for cycle 1 were collected from March 2007 through February 2009 at 15 sites across the country for respondents aged 6 to 79. Full-time members of the Canadian Forces and residents of Crown lands, Indian reserves, institutions and certain remote regions were excluded. The sample represented approximately 96% of the population.15

The CHMS consisted of a household interview during which information about socio-demographic characteristics, health and lifestyle was gathered. This was followed by a visit to a mobile examination centre where direct measurements, including the collection of urine samples, were taken.

Of the households selected for the survey, 69.6% agreed to participate. One or two members of each responding household were invited to take part in the survey. Of these, 88.3% responded to the household questionnaire, and 84.9% of those who completed the questionnaire visited the mobile examination centre. The overall response rate, after adjusting for the sampling strategy, was 51.7%. For adults aged 20 to 79, the overall response rate was 50.9%, and for youth aged 12 to 19, 52.7%. In total, 4,530 respondents aged 12 to 79 participated in the mobile examination centre component of the CHMS.

Ethics approval for conducting the CHMS was obtained from Health Canada’s Research Ethics Board. Written informed consent was obtained from respondents. Participation was voluntary; respondents could opt out of any part of the survey at any time. Additional information about the survey is available in published reports16-20 and on Statistics Canada’s website (http://www.statcan.gc.ca).

During the household interview, respondents were asked if they currently smoked cigarettes daily, occasionally or not at all. They were also asked if they had smoked cigars or a pipe or used snuff or chewing tobacco in the past month. To facilitate accurate reporting, when respondents aged 12 to 19 were being asked about sensitive topics including smoking, parents and guardians were requested to leave the room.

Respondents were asked if they had used prescription or over-the-counter medications in the past month. When they went to the mobile examination centre, they were asked to: confirm the medications they had previously reported; report any other medications they were taking; and report the last time they had taken each medication. Drug Identification Numbers (DIN) were collected for these medications and coded using the Anatomical Therapeutic Chemical (ATC) classification system. ATC code N07BA01 refers to medications in which nicotine is an active ingredient.21 This code would identify smoking cessation aids (nicotine patches, gums and aerosols) that contain nicotine as the active ingredient.

In the introduction to the household interview, respondents were told that direct measurements, including urine samples, would be taken at the mobile examination centre, and were given a list of the laboratory tests that would be performed. However, whether they were aware that the results of the cotinine test could be used to assess smoking status is unknown.

Urinary cotinine analysis
During each respondent’s visit to the mobile examination centre (one day to six weeks after the household interview, an average of 13 days), a spot midstream urine sample was collected in a 120 ml container. The samples were frozen at -20 °C and shipped on dry ice to the testing laboratory at the Institut national de santé publique du Québec (accredited under ISO 17025). Free cotinine was recovered by solid-phase extraction in a 96 well plate format on an automated robotic workstation.22 Deuterated cotinine was used as the internal standard. The extract was then redissolved into 250 μL of mobile phase, and 10 μL were injected into the ultra-performance liquid chromatography-tandem mass spectrometric instrument, operated in the MRM mode with ion source in positive electrospray. The limit of detection was 1.1 μg/L. Details of the quality assurance program at the mobile examination centre and at the laboratory that performed the cotinine testing can be found elsewhere.20

Statistical analysis
Certain exclusions were necessary to compare smoking status based on self-report versus urinary cotinine concentration. Respondents were excluded from the analyses if they:
- did not have a valid cotinine test result, for example, insufficient volume of urine collected; refused urine sample; etc. (n=48).
- reported using a medication with nicotine as an active ingredient (ATC code N07BA01) in the past month (n=4).
- reported smoking cigars or a pipe, or using snuff or chewing tobacco (n=258).

The latter two exclusions were necessary because it is possible that respondents who reported not being cigarette smokers could have been classified as smokers based on elevated cotinine concentrations that resulted from using these other nicotine-containing products. Among the 4,530 CHMS mobile examination centre participants aged 12 to 79, these exclusions resulted in a loss of 307 cases, leaving a final sample size of 4,223 for the study. (Three records were flagged for exclusion for more than one reason.)

For smoking status based on self-report, respondents who reported that they currently smoked cigarettes “daily” or “occasionally” were classified as smokers. For smoking status based on cotinine concentrations, respondents with urinary concentrations greater than 50 ng/ml were classified as smokers. This is the cut-point recommended by the Society for Research on Nicotine and Tobacco to distinguish tobacco users from non-tobacco users, including those
exposed to second-hand smoke.\textsuperscript{23} It is highly unlikely that levels above this cut-point would be observed among non-users, even if they were regularly exposed to second-hand smoke.\textsuperscript{23,24}

The correlation between smoking prevalence based on self-report and cotinine concentrations was calculated. The accuracy of self-reported smoking status was assessed by calculating sensitivity and specificity. Sensitivity is the percentage of true positives (the percentage of respondents who reported being smokers among those classified as smokers based on cotinine concentrations). Specificity is the percent of true negatives (the percentage of respondents who reported being non-smokers among those classified as non-smokers based on cotinine concentrations).

Comparisons were made between the self-reported prevalence of smoking based on the CHMS and on other Statistics Canada surveys that collect data on smoking status. To make meaningful comparisons, it was necessary to calculate smoking prevalence for the entire CHMS sample (\(n=4,530\)) without the exclusions in the sensitivity and specificity analyses. CHMS smoking prevalence estimates based on the entire sample are shown in Appendix Table A.

Results

Smoking prevalence: Self-report versus urinary cotinine

According to the CHMS, the prevalence of smoking was 18.8% based on self-report and 19.1% based on urinary cotinine concentration (Table 1). Differences between prevalences based on self-report versus cotinine concentration were not significant for any of the age/sex groups. Correlation results indicated strong agreement between smoking status based on self-report and cotinine (\(r=0.90, p<0.001\)).

### Table 1

<table>
<thead>
<tr>
<th>Sex/Age group (years)</th>
<th>Self-report</th>
<th></th>
<th>Urinary cotinine concentration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>95% confidence interval</td>
<td>%</td>
<td>95% confidence interval</td>
</tr>
<tr>
<td>Total</td>
<td>18.8</td>
<td>16.3</td>
<td>21.5</td>
<td>19.1</td>
</tr>
<tr>
<td>12 to 19</td>
<td>7.7\textsuperscript{e}</td>
<td>4.4</td>
<td>13.0</td>
<td>5.8\textsuperscript{e}</td>
</tr>
<tr>
<td>20 to 79</td>
<td>20.2</td>
<td>17.8</td>
<td>23.0</td>
<td>20.8</td>
</tr>
<tr>
<td>Male</td>
<td>19.0</td>
<td>15.8</td>
<td>22.7</td>
<td>19.2</td>
</tr>
<tr>
<td>12 to 19</td>
<td>5.4\textsuperscript{e}</td>
<td>3.1</td>
<td>9.4</td>
<td>4.5\textsuperscript{e}</td>
</tr>
<tr>
<td>20 to 79</td>
<td>20.9</td>
<td>17.5</td>
<td>24.7</td>
<td>21.2</td>
</tr>
<tr>
<td>Female</td>
<td>18.6</td>
<td>15.6</td>
<td>22.1</td>
<td>19.0</td>
</tr>
<tr>
<td>12 to 19</td>
<td>10.0\textsuperscript{e}</td>
<td>6.2</td>
<td>18.5</td>
<td>7.1\textsuperscript{e}</td>
</tr>
<tr>
<td>20 to 79</td>
<td>19.7</td>
<td>16.6</td>
<td>23.2</td>
<td>20.4</td>
</tr>
</tbody>
</table>

\textsuperscript{e} use with caution

Note: Excludes respondents who did not have a valid cotinine test result, reported using a medication with nicotine as an active ingredient in the past month, or reported smoking cigars or a pipe, or using snuff of chewing tobacco.


### Table 2

<table>
<thead>
<tr>
<th>Urinary cotinine concentration</th>
<th>Self-reported smoker</th>
<th></th>
<th>Self-reported non-smoker</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>95% confidence interval</td>
<td>%</td>
<td>95% confidence interval</td>
</tr>
<tr>
<td>Less than or equal to 50 ng/ml (non-smoker)</td>
<td>1.7</td>
<td>1.1</td>
<td>2.6</td>
<td>98.3</td>
</tr>
<tr>
<td>More than 50 ng/ml (smoker)</td>
<td>91.6</td>
<td>86.3</td>
<td>95.0</td>
<td>8.4</td>
</tr>
</tbody>
</table>

Note: Excludes respondents who did not have a valid cotinine test result, reported using a medication with nicotine as an active ingredient in the past month, or reported smoking cigars or a pipe, or using snuff of chewing tobacco.


Accuracy of self-reported smoking status

Sensitivity for self-reported smoking status was 91.6% (Tables 2 and 3). That is, among respondents classified as smokers based on their urinary cotinine concentration, 91.6% reported that they were cigarette smokers, and 8.4% were misclassified in that they reported that they did not smoke cigarettes. The mean cotinine concentration for the misclassified cases was 615.7 ng/ml (95% CI: 427.5 to 803.8), which was substantially lower than the mean (1,239.4 [95% CI: 1100.2 to 1378.7]) for properly classified cases. Among the
Specificity

classiﬁcation status was 98.3%, meaning that 1.7% of they usually smoked only one cigarette smoked on 10 or fewer days in the past these occasional smokers said they had were occasional smokers. Most (82%) of report. Urinary cotinine concentration prevalence in Canada based on self-validity of estimates of cigarette smoking was 0.6 percentage points lower when United States where smoking prevalence was 0.3 percentage points lower based on self-report closely approximates estimates based on cotinine concentration.

Smoking prevalence was 0.3 percentage points lower based on self-report than on cotinine concentrations. This was consistent with results from the United States where smoking prevalence was 0.6 percentage points lower when based on self-report than on cotinine concentration.14 In England and Poland, smoking prevalence based on self-report was lower by 2.8 percentage points and 4.4 percentage points, respectively.14 The strong correlation and lack of significant differences between smoking prevalence based on self-report and cotinine concentration in the present study suggest that self-reported data provide a valid estimate of national smoking prevalence in Canada.

Although sensitivity was high (91.6%), 8.4% of respondents were classified as “false negatives” (their cotinine concentrations identiﬁed them as smokers although they reported that they did not smoke). The mean cotinine concentration was substantially lower among these false negatives (615.7 ng/ml) than among properly classiﬁed cases (1239.4 ng/ml), suggesting that heavy smokers are more likely than light smokers to report that they smoke.

Some misreporting would be expected due to social desirability bias. Although parents/guardians were asked to leave the room when the questions on smoking were administered to respondents aged 12 to 19, some of these younger respondents may have been reluctant to report that they smoked, resulting in the lower sensitivity estimates for this age group.

However, other reasons may explain some of the false negatives. Consistent with previous research,29 a signiﬁcantly higher percentage of the false negative cases reported being former smokers rather than never smokers, and the majority of these former smokers were recent quitters. Relapse is common among recent quitters.29 If some of them relapsed in the period between their household interview and mobile examination centre visit, they would have been inappropriately classiﬁed as false negatives. Similarly, smoking initiation or experimentation in this period may have resulted in some cases being inappropriately classiﬁed as false negatives, particularly among respondents aged 12 to 19.

Other studies have found varying levels of sensitivity for self-reported estimates of smoking, depending on the population studied, the type of biological specimen used in the measurement of cotinine, and the cut-points used to identify smokers.13 Similar to the CHMS ﬁndings, sensitivity estimates greater than 90% have frequently been reported,13 but studies based on pregnant women,30 and on patients with smoking-related illnesses such as respiratory disease31 and cancer32 have yielded lower estimates of sensitivity.

A small percentage (1.7%) of respondents were classiﬁed as “false positives” (their cotinine concentration classiﬁed them as non-smokers, but they reported that they smoked cigarettes). Nearly all these false positive cases reported that they were occasional

Table 3
Sensitivity and speciﬁcity of self-reported smoking status, by sex and age group, household population aged 12 to 79, Canada, March 2007 to February 2009

<table>
<thead>
<tr>
<th>Sex/Age group (years)</th>
<th>Sensitivity %</th>
<th>95% conﬁdence interval</th>
<th>Specificity %</th>
<th>95% conﬁdence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>91.6</td>
<td>96.3 to 95.0</td>
<td>98.3</td>
<td>97.4 to 98.9</td>
</tr>
<tr>
<td>12 to 19</td>
<td>81.6</td>
<td>95.5 to 95.4</td>
<td>96.9</td>
<td>94.0 to 96.4</td>
</tr>
<tr>
<td>20 to 79</td>
<td>92.0</td>
<td>96.3 to 95.4</td>
<td>98.6</td>
<td>97.8 to 99.1</td>
</tr>
<tr>
<td>Male</td>
<td>92.1</td>
<td>96.7 to 95.4</td>
<td>98.4</td>
<td>97.3 to 99.1</td>
</tr>
<tr>
<td>12 to 19</td>
<td>76.3</td>
<td>93.0 to 95.5</td>
<td>97.9</td>
<td>94.7 to 99.2</td>
</tr>
<tr>
<td>20 to 79</td>
<td>92.6</td>
<td>97.4 to 95.7</td>
<td>98.5</td>
<td>97.4 to 99.1</td>
</tr>
<tr>
<td>Female</td>
<td>91.2</td>
<td>84.4 to 95.1</td>
<td>98.3</td>
<td>96.9 to 99.0</td>
</tr>
<tr>
<td>12 to 19</td>
<td>85.0</td>
<td>93.3 to 96.6</td>
<td>95.8</td>
<td>92.6 to 98.3</td>
</tr>
<tr>
<td>20 to 79</td>
<td>91.4</td>
<td>83.9 to 95.6</td>
<td>98.6</td>
<td>97.6 to 99.2</td>
</tr>
</tbody>
</table>

Note: Excludes respondents who did not have a valid cotinine test result, reported using a medication with nicotine as an active ingredient in the past month, or reported smoking cigars or a pipe, or using snuff of chewing tobacco.


Discussion

This study is the ﬁrst to examine the validity of estimates of cigarette smoking prevalence in Canada based on self-report. Urinary cotinine concentration measured by the CHMS provided a biomarker of tobacco smoke exposure with which to validate self-reported smoking status. The results indicated that smoking prevalence based on self-report closely approximates estimates based on cotinine concentration.

Sensitivity estimates were similar for males and females. And although sensitivity tended to be lower for respondents aged 12 to 19 than for those aged 20 to 79, the difference did not attain statistical signiﬁcance.

Speciﬁcity for self-reported smoking status was 98.3%, meaning that 1.7% of respondents whose cotinine concentration classiﬁed them as non-smokers reported that they smoked cigarettes. Of these, the majority (89%) reported that they were occasional smokers. Most (82%) of these occasional smokers said they had smoked on 10 or fewer days in the past month, and about half (51%) reported they usually smoked only one cigarette on the days that they smoked.
smokers, and most reported smoking on 10 or fewer days in the past month. Cotinine is a measure of recent exposure to tobacco smoke, so it is likely that cotinine levels in some of these occasional smokers were too low to classify them as smokers.

An important question is the degree to which findings from this study apply to other Statistics Canada surveys that collect self-reported smoking data, such as the Canadian Community Health Survey (CCHS) and the Canadian Tobacco Use Monitoring Survey (CTUMS). Survey respondents may be more likely to accurately report their smoking status if they know, or believe, that a biospecimen will also be collected to determine smoking status. Unlike the CHMS, biospecimens are not collected by the CCHS or CTUMS.

According to the CHMS, 20% of Canadians aged 12 to 79 were self-reported smokers (Appendix Table A). The prevalence of self-reported smoking among people aged 12 or older from the 2009 CCHS was 20%. These similar results suggest that self-reported CCHS data provide accurate estimates of cigarette smoking prevalence.

In 2009, the prevalence of smoking estimated from CTUMS was 18% among the population aged 15 or older. While trends in CTUMS data paralleled those derived from the CCHS, CTUMS smoking rates were consistently lower. However, unlike CTUMS, which is designed to monitor smoking prevalence, the smoking questions in the CCHS (and the CHMS) were asked in the context of a general health survey. A study of why smoking prevalence differs between the CCHS and CTUMS suggested that people are more inclined to talk frankly about smoking when the topic is part of a broad-based health survey.

The way in which data were collected might also contribute to differences in prevalence estimates between the surveys. CTUMS is conducted entirely by telephone; the CHMS is conducted entirely in person; and the CCHS uses in-person and telephone interviews. Nonetheless, a study comparing the effect of in-person and telephone interviews found that, overall, the interview mode was not associated with significantly different estimates of smoking prevalence.

**Limitations**

One limitation of this study was the relatively low overall CHMS response rate (52%). While the survey weights ensured that the sample was representative of the target population, bias might exist if non-respondents were more or less likely than respondents to be cigarette smokers and/or more or less likely to accurately self-report their smoking status. However, a comparison of the characteristics of those who responded to the household questionnaire with the characteristics of people who went on to complete the mobile examination centre component found the prevalence of smoking to be similar in the two groups. Furthermore, smoking prevalence based on self-report was similar in the CHMS and the CCHS, the latter of which had a higher response rate (73%).

The examination centre visit occurred, on average, 13 days after the household interview. Although the number of cases would likely be small, a true change in smoking behaviour during this interval might have resulted in some respondents being erroneously classified as false negatives or false positives.

The use of cotinine concentrations to assess the validity of self-reported smoking status may be inappropriate for occasional smokers and result in some respondents being erroneously classified as false positives.

Small sample sizes for respondents aged 12 to 19 resulted in estimates with high sampling variability. Therefore, results for this age group should be interpreted with caution. Possibly because of small sample sizes, the lower sensitivity for the younger age group did not attain statistical significance. As future CHMS cycles become available, it will be possible to augment the sample and produce estimates with higher reliability.

**Conclusion**

Representative data for the Canadian population showed no significant difference between national estimates of smoking prevalence based on self-report versus urinary cotinine concentration. This suggests that self-reported data on smoking status provide a valid estimate of the prevalence of smoking in Canada.

**Acknowledgments**

The authors thank Tracey Bushnik and Johanne Levesque of Statistics Canada for their contributions to this study. Scott Leatherdale is a Cancer Care Ontario Research Chair in Population Studies.
References


Table A
Prevalence of cigarette smoking based on self-report, by age group and sex, household population aged 12 to 79, Canada, March 2007 to February 2009

<table>
<thead>
<tr>
<th>Age group</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% confidence interval</td>
<td>95% confidence interval</td>
<td>95% confidence interval</td>
</tr>
<tr>
<td></td>
<td>% from to</td>
<td>% from to</td>
<td>% from to</td>
</tr>
<tr>
<td>Total</td>
<td>20.3 17.8 23.1</td>
<td>21.8 19.0 24.8</td>
<td>18.9 15.8 22.5</td>
</tr>
<tr>
<td>12 to 18</td>
<td>11.7† 6.3 20.7</td>
<td>10.9† 6.2 18.4</td>
<td>12.6† 6.1 24.2</td>
</tr>
<tr>
<td>20 to 79</td>
<td>21.5 19.4 23.9</td>
<td>23.3 20.7 26.2</td>
<td>19.8 16.9 23.1</td>
</tr>
</tbody>
</table>

† use with caution