

How Does the Exclusion of Several Population Groups from the National Health Interview Survey Affect Estimates of US Cigarette Smoking Prevalence?

Ken Warner, University of Michigan School of Public Health

Co-authors: Ruoyan Sun, UAB, and David Mendez, UM

Disclosures

This research was funded by the National Cancer Institute (NCI) and Food and Drug Administration (FDA) grant U54CA229974.

	Tobacco Industry	Non-Pharma Nicotine Product Industry	Pharma Industry
The work being presented has received funding or other means of support from any of the following sources:	No	No	No
Any of the authors have received funding (including consultancy) from any of the following sources in the past 5 years:	No	No	No

The authors have no conflicts of interest.

BREAKING NEWS!

US adult smoking prevalence in 2024: **9.9%!**

CDC: Early Release from the 2024 National Health Interview Survey (NHIS), June 2025

But is it so?

Context: NHIS designed to be representative of the non-institutionalized civilian population. Sampling frame does not include:

- Active-duty military
- Institutionalized people in facilities like nursing homes, prisons and jails and psychiatric hospitals
- Hard-to-reach groups, like unsheltered homeless people



UK study

How much does the absence of the 'hidden population' from United Kingdom household surveys underestimate smoking prevalence?

Emma Beard^{1,2,3} | Lion Shahab^{1,3} | Jamie Brown^{1,3} | Sharon Cox^{1,3}

Examined how much exclusion of groups from the UK Annual Population Survey affected APS estimate of smoking prevalence.

- APS estimate for 2022 = 12.9%
- Adjusted estimate = 14.5-14.8%, 12-14% higher than APS

Groups not covered by NHIS

1. Active-duty military.
2. Long-term care facilities, including nursing homes, assisted living, and residential senior-care.
3. Group homes for ≥ 7 people with intellectual and developmental disabilities (IDD). (Group homes of < 7 people with IDD should be covered by NHIS.)
4. Psychiatric inpatient hospitals/units.
5. Residential substance-use treatment programs.
6. Federal, state, and local prisons and jails.
7. Immigration detention centers.
8. Unsheltered homeless. (NHIS includes people in homeless shelters.)
9. Sofa (couch) surfers.

Methods

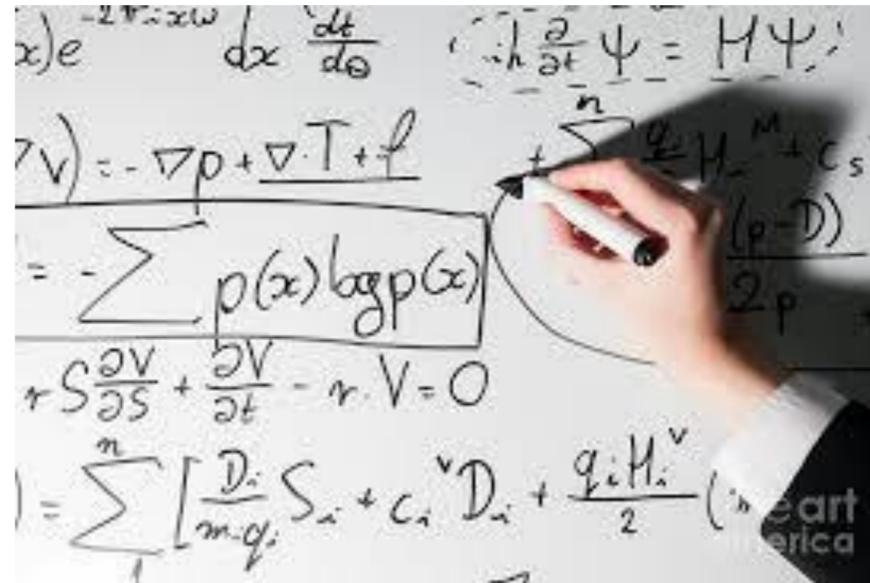
1. Estimate size of population and smoking prevalence in 2024 for each group.
 - a. Some precise (e.g., population of active-duty military; prisoners)
 - b. Some good existing estimates (e.g., population and smoking rate of unhoused homeless)
 - c. Several very imprecise; require estimating from multiple sources (e.g., population of sofa surfers)
 - d. Sources for all estimates: government reports, peer-reviewed studies, reports from nonprofit organizations (See 12-page supplement on sources)
 - e. Develop lower/upper bounds around best estimates (95% CI's when available)

Methods (cont'd.)

2. Assuming a triangular or normal (when CIs available) distribution, we sampled the ranges of population size and smoking prevalence to obtain point estimates.
3. Multiplied population x smoking prevalence to estimate number of people who smoked in each group.
4. Summed across all groups to estimate population size and number of smoking adults not covered by NHIS.
5. To estimate number of smoking adults eligible for inclusion in NHIS, we subtracted the study groups' population from 2024 adult population (267 million) and multiplied by NHIS's estimated 2024 smoking prevalence (9.9%).

Methods (cont'd.)

6. Combined total number of smoking adults in NHIS-included and omitted groups and divided by the total adult population to estimate national smoking prevalence including populations not sampled by NHIS.
7. Conducted a Monte Carlo simulation with 100,000 replications to generate stable point estimates and confidence intervals.



Results

Note: Following results differ from those in abstract. These are final results.

Table 1. Population and smoking prevalence estimates for each group omitted from NHIS and the resulting number of people who smoke

Group	Best population estimate (1)	Best smoking estimate (2)	No. who smoke (3) = (1) x (2)
Active-duty military	1,270,000	18.4%	233,680
Long-term care	2,400,000	5.3%	127,200
IDD	85,800	18.5%	15,873
Psychiatric inpatient	120,000	39.6%	47,520
SUD inpatient	155,000	59.7%	92,535
Prisons and jails	1,840,000	60.0%	1,104,000
Immigration detention	37,500	7.8%	2,925
Unsheltered homeless	274,000	74.7%	204,678
Sofa surfers	2,700,000	48.7%	1,314,900
TOTAL	8,882,300	35.4%	3,143,311

Results (cont'd.)

Summary of table 1

1. Individual groups' population sizes ranged from 37,500 (immigration detention) to 2.7 million (sofa surfers).
2. Total population not included in NHIS's sampling frame = 8.88 million people.
3. Smoking prevalence ranged from 5.3% (people in long-term care) to 74.7% (unsheltered homeless people).
4. With 3.14 million smoking, population-weighted smoking prevalence of all 9 groups combined was 35.4%.

Results (cont'd.)

Table 2. Results of Monte Carlo analysis

Output measures	Mean	95% confidence interval
Population excluded from NHIS sampling frame (millions)	8.88	8.41, 9.35
Individuals from excluded groups who smoke (millions)	3.14	2.76, 3.55
Smoking prevalence of individuals from excluded groups	35.4%	31.5%, 39.3%
Population included in NHIS sampling frame (millions)	258.1	257.7, 258.6
Individuals from included groups who smoke (millions)	25.6	24.4, 26.7
Smoking prevalence of individuals from included groups	9.9%	9.4%, 10.3%
Overall smoking prevalence (including NHIS-omitted groups)	10.75%	10.29%, 11.20%
Overall smoking prevalence percentage above 9.9%	8.6%	4.0%, 13.2%

Results (cont'd.)

Summary of table 2

1. Number of people who smoke in the populations omitted and included by NHIS were, respectively, 3.14 million and 25.6 million, for a total of 28.7 million.
2. The omitted smoking population constituted 10.9% of all US adults who smoked in 2024.
3. The omitted population comprised 3.3% of total US adult population.
4. The mean of US smoking prevalence, including the NHIS-excluded population, was 10.75%, 8.6% higher than the standard national estimate of 9.9%.

Limitations

Several population and smoking prevalence estimates, while data informed, were necessarily somewhat arbitrary. For those with the least certainty, we included a wide uncertainty interval around our best estimates.

Conclusions

1. The smoking prevalence of groups omitted from NHIS is >3 times that of those included (35.4% vs. 9.9%). As a result, their omission from NHIS leads to underestimation of national smoking prevalence.
2. Importantly, the amount of underestimation will grow over time
 - a. if the NHIS-omitted groups continue to decrease their smoking much more slowly than the NHIS-covered population; and/or
 - b. If the omitted groups' populations grow more than does the general population.

Conclusions (cont'd.)

3. To accurately characterize Americans' smoking behavior, future NHIS estimates of smoking prevalence should emphasize their underestimation of true US adult smoking.
4. Explaining why will emphasize that the burden of smoking is borne disproportionately by a number of marginalized groups, explicitly including those not covered by NHIS.
5. It will also focus attention on the need for smoking cessation efforts targeting some of the nation's highest-smoking marginalized groups.

BREAKING NEWS!

Did we cross the magical 10% smoking prevalence barrier in 2024?

Not quite.

THANK YOU!

