Novel microsimulation model of tobacco use behaviours and outcomes:

calibration and validation in a US population

Supplement

Krishna P. Reddy, Alexander J.B. Bulteel, Douglas E. Levy, Pamela Torola, Emily P. Hyle, Taige Hou, Benjamin Osher, Liyang Yu, Fatma M. Shebl, A. David Paltiel, Kenneth A. Freedberg, Milton C. Weinstein, Nancy A. Rigotti, Rochelle P. Walensky

Table of Contents

METHODS: SUPPLEMENTARY TEXT

National Health Interview Survey (NHIS)

NHIS is a yearly in-person questionnaire administered to the civilian noninstitutionalized population of the US which, since 1965, has collected information on individual smoking status. Since 1991, the smoking section of the NHIS has first queried "ever smoker" status – defined as having smoked at least 100 cigarettes in one's lifetime – then asked ever smokers, "Do you NOW smoke cigarettes every day, some days or not at all?" which resulted in the classification of occasional smokers as current smokers, even though they initially may have said they did not smoke now. Regardless of response to the second question, participants were asked about the frequency of their smoking.[1]

Use of NHIS data in the Simulation of Tobacco and Nicotine Outcomes and Policy (STOP) model

We downloaded NHIS and linked National Death Index (NDI) data from the Integrated Public Use Microdata Series (IPUMS)Health Surveys for the years 1997-2009.[2] We obtained basic demographic information, smoking status and behavioral variables, and death status/year of death reported through 2011. For all derivations below, we excluded those with unknown smoking or mortality status. Only people 18 years of age or older were surveyed about tobacco use, and since ages 85+ years are censored, we excluded those as well.

External validation - mortality

From the pooled 1997-2009 data, we used the IPUMS-recoded and -constructed survey weights adjusting for differential representation in the smoking sub-sample and NDI follow-up. Because of the NHIS sampling design, weights must be used so that the survey respondents can be collectively expanded to represent the civilian noninstitutionalized population of the US. These weights represent a

surveyed individual's inverse probability of being included in both the survey supplement, which containsquestions about smoking, and the NHIS-NDI linked mortality files.

Smoking cessation inputs: We derived age- and sex-stratified cessation rates using the NHIS variable that reported years since respondents (former smokers) quit smoking. We excluded quit ages before age 16 years (due to perceived inconsistency of coding of the "time since quit" variable – some entries implied negative quit age) and included quit ages through age 85 years.

Smoking initiation inputs: Similarly, we derived initiation rates, also age- and sex-stratified, using the variable that reported years since the respondent (a current or former smoker) started to smoke regularly. We used initiation rates from ages 6 to 61 years, the last age for which data are consecutively available for both women and men.

Mortality inputs: We calculated mortality rates by age, sex, smoking status, and five-year age group using the smoking status variable and NDI reporting through 2011. The rates were calculated from a follow-up period beginning in the respondent's survey year and ending in 2011 or the year of her/his death, whichever came first. Because smoking behavior data were collected only at baseline in NHIS and not again at the time of death, there may have been some misclassification of smoking status (e.g., someone who was a current smoker at the time of NHIS assessment may have subsequently quit and later died but was still considered a current smoker). However, all-cause mortality rates do not significantly decrease until a few years after cessation, and we considered those who had quit smoking to have similar mortality risks to current smokers until five years of abstinence. This reflects contemporary studies in which former smokers were defined as those who had not smoked in the previous five years and data

from large US cohort studies that indicate that the all-cause mortality risk in men who quit smoking does not fall below that of current smokers until five years of abstinence.[3,4]

Input cohort: We derived initial cohort characteristics for this analysis from theNHIS 1997-2009 pooled dataset: current, former and never smoking prevalence by five-year age group from age <20 to 84 years; mean (standard deviation [SD]) age adjusted for bounding of the distribution; sex distribution; and mean (SD) time since quit bounded by a minimum quit age of 16 years.

Output comparison: We compared STOP model output for mortality to NHIS/NDI-reported mortality in the form of mortality rates and cumulative mortality from age 20 years.

External validation – smoking prevalence

From the NHIS 1997 data, using the same survey weights as above:

Smoking cessation, smoking initiation, and mortality inputs: We used the same cessation, initiation, and mortality rate inputs as for the mortality validation.

Input cohort: We derived input cohort characteristics for this analysis from the 1997 NHIS data: current, former, and never smoking prevalence by five-year age group from age <20 to 84 years; mean (SD) age adjusted for bounding of the distribution; sex distribution; and mean (SD) time since quit bounded by a minimum quit age of 16 years.

Output comparison: Starting with the 1997 NHIS respondents and following themeach year from 1998 to 2009, we compared STOP projections to NHIS data regarding prevalenceof current, former, and never smokers by five-year age group from age 30 to 84 years.

REFERENCES IN SUPPLEMENT

1 Centers for Disease Control and Prevention. National Health Interview Survey. Adult tobacco use - major changes. https://www.cdc.gov/nchs/nhis/tobacco/tobacco_changes.htm (accessed 9 Jan 2020).

2 Blewett LA, Rivera Drew JA, Griffin R, *et al.IPUMS Health Surveys: National Health Interview Survey, Version 6.3 [dataset]*. Minneapolis, MN: IPUMS. https://nhis.ipums.org (accessed 9 Jan 2020).

3 Thun MJ, Carter BD, Feskanich D, *et al.* 50-year trends in smoking-related mortality in the United States. *N Engl J Med* 2013;**368**:351–64. doi:10.1056/NEJMsa1211127

4 Jha P, Ramasundarahettige C, Landsman V, *et al.* 21st-century hazards of smoking and benefits of cessation in the United States. *N Engl J Med* 2013;**368**:341–50. doi:10.1056/NEJMsa1211128

5 Tonstad S, Tønnesen P, Hajek P, *et al.* Effect of maintenance therapy with varenicline on smoking cessation: a randomized controlled trial. *JAMA* 2006;**296**:64–71. doi:10.1001/jama.296.1.64

6 Brendryen H, Kraft P. Happy Ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. *Addiction* 2008;**103**:478–84. doi:10.1111/j.1360-0443.2007.02119.x

7 Bullen C, Howe C, Laugesen M, *et al.* Electronic cigarettes for smoking cessation: a randomised controlled trial. *Lancet* 2013;**382**:1629–37. doi:10.1016/S0140-6736(13)61842-5

8 Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction* 2004;**99**:29–38. doi:10.1111/j.1360-0443.2004.00540.x

Supplementary Table 1. Examples of transition probabilities in the simulation model.

Parameter	Value	Reference
Smoking initiation probability, monthly, by age and sex	0-0.0063	Derived from NHIS 1997-2009 data
Cessation probability, monthly, by age and sex	0-0.035	Derived from NHIS 1997-2009 data
Relapse probability, monthly, by time since cessation	$P_{Relapse}= 0.62*e^{-0.33*t}$	[5–8]
Never Smokermortality probability, monthly, by age and sex, x 10^{-4}	0.4-95.2	Derived from NHIS/NDI 1997-2009 data
Current Smoker or Recent Quitter mortality probability, monthly, by age and sex, x 10^{-4}	0.4-136.1	Derived from NHIS/NDI 1997-2009 data
Former Smoker mortality probability, monthly, by age, sex, and time since cessation[a], x 10^{-4}	0-111.5	Derived from NHIS/NDI 1997-2009 data
Time to transition from Recent Quitter to Former Smoker	5 years[b]	Assumption, [3,4]

NHIS: National Health Interview Survey. NDI: National Death Index. STOP: Simulation of Tobacco and Nicotine Outcomes and Policy.

[a]These former smoker mortality probabilities were used only in the external validation exercises, comparing STOP model output to NHIS/NDI results. In cross-validation comparing STOP model output to

the CISNET model results, we applied multipliers to never smoker mortality rates to derive former smoker mortality rates and then converted these to probabilities.

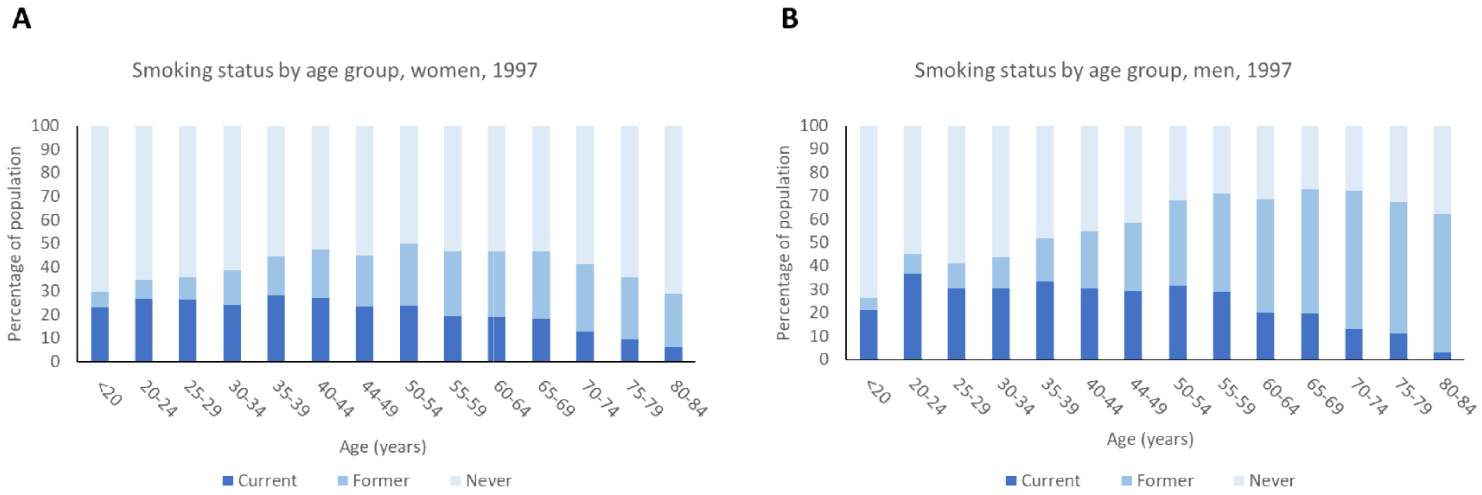
[b]This five-year abstinence period was based on data showing mortality risks by years since smoking cessation.[3,4] Weperformed an analysis in which we assumed that mortality risks decreased to "Former Smoker" levels immediately upon quitting smoking, but the model-generated results were not a better fit to NHIS data (results not shown).

Supplementary Table 2. **Selected age- and sex-specific annual mortality rates: STOP model output versus NHIS-derived mortality.**

Women						
	Current Smoker		Former Smoker		Never Smoker	
Age group, years	STOP	NHIS/NDI	STOP	NHIS/NDI	STOP	NHIS/NDI
40-44	0.0023	0.0023	0.0017	0.0015	0.0012	0.0012
50-54	0.0071	0.0071	0.0035	0.0029	0.0025	0.0025
60-64	0.0147	0.0146	0.0087	0.0079	0.0051	0.0051
Men						
	Current Smoker		Former Smoker		Never Smoker	
Age group, years	STOP	NHIS/NDI	STOP	NHIS/NDI	STOP	NHIS/NDI
40-44	0.0031	0.0034	0.0028	0.0025	0.0018	0.0018
50-54	0.0094	0.0096	0.0059	0.0053	0.0042	0.0043
60-64	0.0233	0.0232	0.0114	0.0099	0.0067	0.0066

STOP: Simulation of Tobacco and Nicotine Outcomes and Policy. NHIS: National Health Interview Survey. NDI: National Death Index.

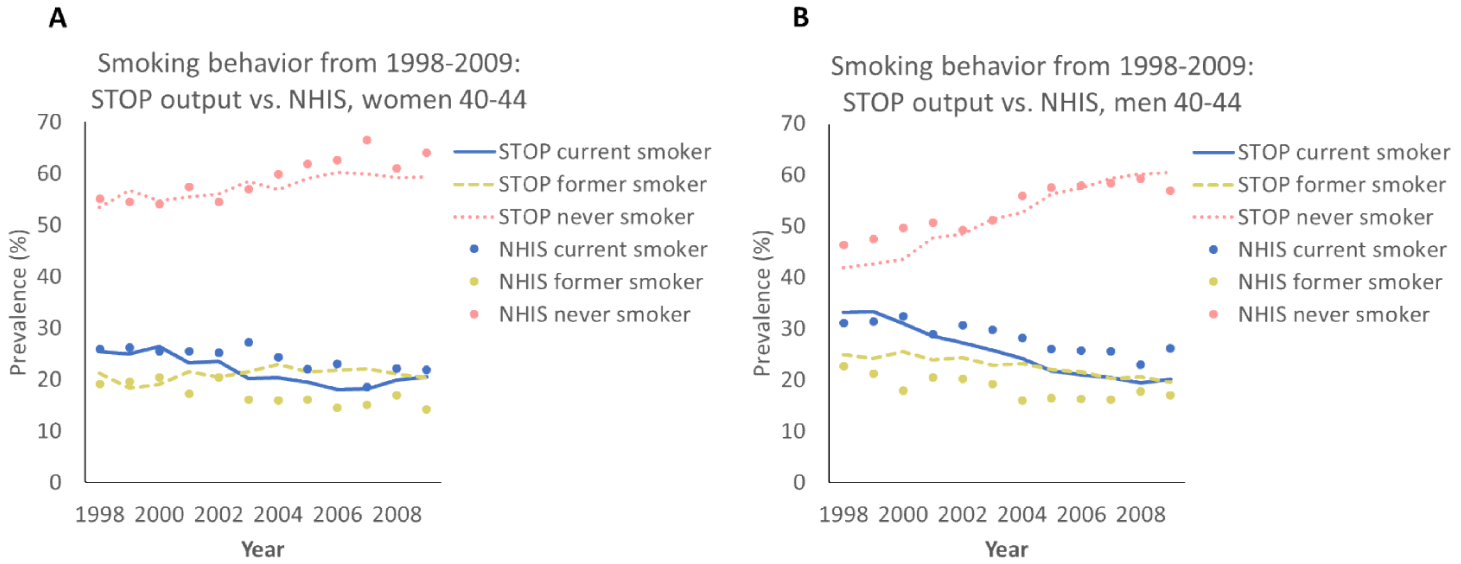
Supplementary Figure 1. Smoking status by age group and sex in 1997: STOP model inputs for the external validation exercise of smoking prevalence.



STOP: Simulation of Tobacco and Nicotine Outcomes and Policy. NHIS: National Health Interview Survey.

Panel A depicts women; Panel B depicts men. For each age group, the prevalence of current, former, and never smokers is entered as STOP model input. The external validation analysis of smoking prevalence, using weighted NHIS 1997 survey respondent data as inputs, projects smoking prevalence annually from 1998 to 2009.

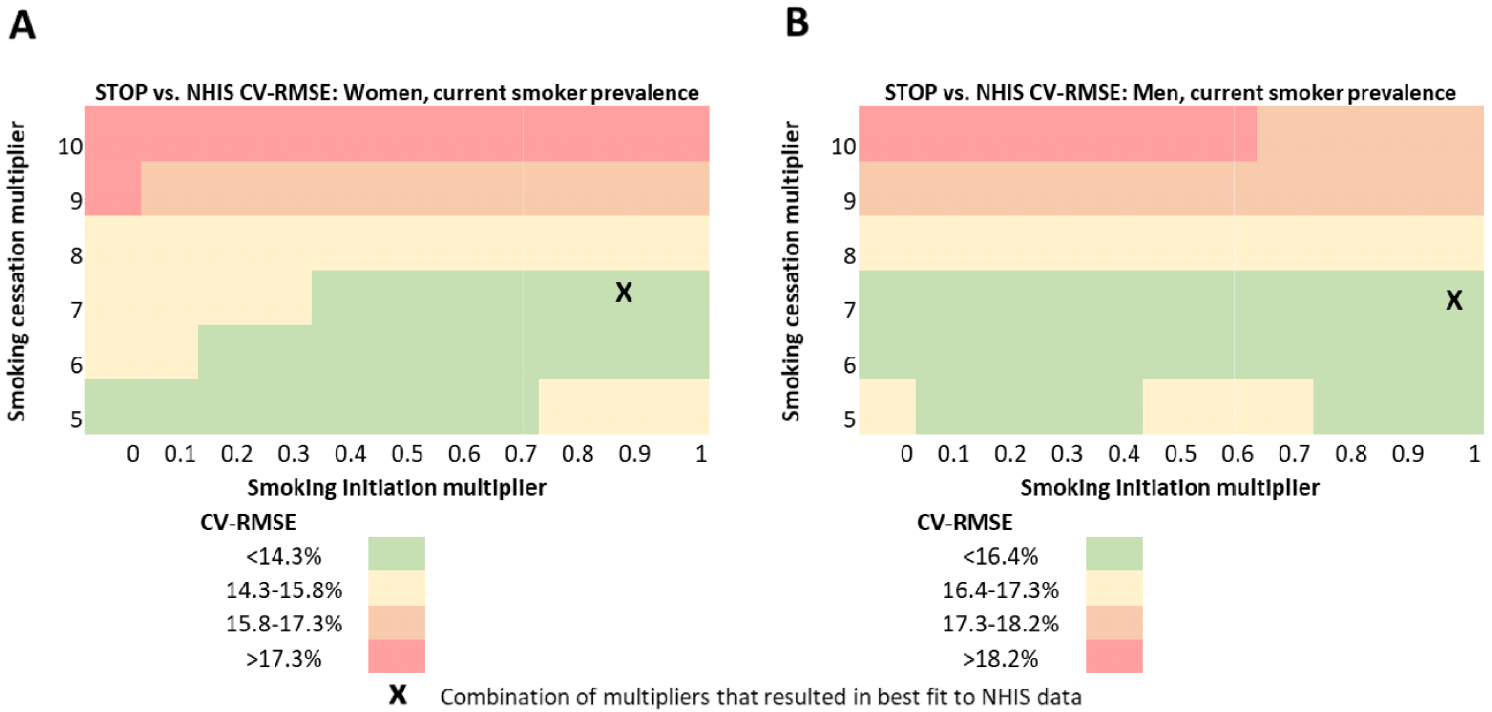
Supplementary Figure 2. External validation: STOP model results and NHIS results for prevalence of current, former, and never smokers each year from 1998 to 2009.



STOP: Simulation ofTobacco and Nicotine Outcomes and Policy. NHIS: National Health Interview Survey. NDI: National Death Index.

Panel A depicts women, Panel B depicts men. The STOP model cohort had the age and smoking status distribution of the 1997 NHIS survey respondents (Supplementary Figure 1). The STOP-generated results reflect a microsimulation that included probabilities of smoking initiation, smoking cessation, smoking relapse, and mortality.

Supplementary Figure 3. Two-way sensitivity analysis of smoking initiation and smoking cessation multipliers for external validation exercise.



STOP: Simulation of Tobacco and Nicotine Outcomes and Policy. NHIS: National Health Interview Survey. CV-RMSE: coefficient of variation of root-mean-square error.

Panel A depicts women; Panel B depicts men. The multipliers were applied to the original cessation and initiation rates derived from pooled 1997-2009 NHIS data and used in the STOP model. The horizontal axis shows the multiplier applied to smoking initiation rates (subsequently converted to probabilities) in the STOP model, and the vertical axis shows the multiplier applied to smoking cessation rates (subsequently converted to probabilities). Colored cells represent the CV-RMSE of STOP model-generated versus NHIS-reported current smoking prevalence among people ages 30-84 yearsfrom 1998 to 2009.