Calibrating Probabilities of Transition in Use of Combustible Cigarettes and E-Cigarettes with Multi-state Modeling

School of Public Health
BROWN UNIVERSITY

Breanna Richards[1]; Thesis Advisor: Stavroula Chrysanthopoulou[1]
[1]Brown University School of Public Health Department of Biostatistics, Providence, RI,

Overview

We utilized data from a nationally representative longitudinal study in the estimation of **behavioral transitions** and **inference** for **tobacco use** in the U.S. These estimates were then used to **validate a microsimulation model** aimed at accurately projecting downstream prevalence of product use.

Background and Study Design

- **Tobacco use** is the leading cause of preventable disease, disability, and death in the United States.
- The U.S. witnessed a **dramatic increase** in the use of electronic nicotine delivery systems (ENDS)/e-cigarettes in the last decade. Most notably, the year 2019 observed a boom in the use of JUUL among high-school-aged youth.
- **Simulation modeling** provides a useful approach to addressing and understanding how **traditional cigarettes and e-cigarettes** interact and affect the prevalence of tobacco use within our ecosystem.
- We fit a **Markov multi-state model** (MMSM) for participants in Waves 1-5 of the Population Assessment of Tobacco and Health (PATH) longitudinal study in order to inform and update the **Simulation of Tobacco and Nicotine Outcomes and Policy** (STOP) model, a microsimulation model used to project effects of tobacco use and cessation over time, including the key behavior of relapse.
- Results from the MMSM were supplemented with results from **mixed-effects regression** which allowed for the quantitative inference of the effect of covariates on use behaviors.

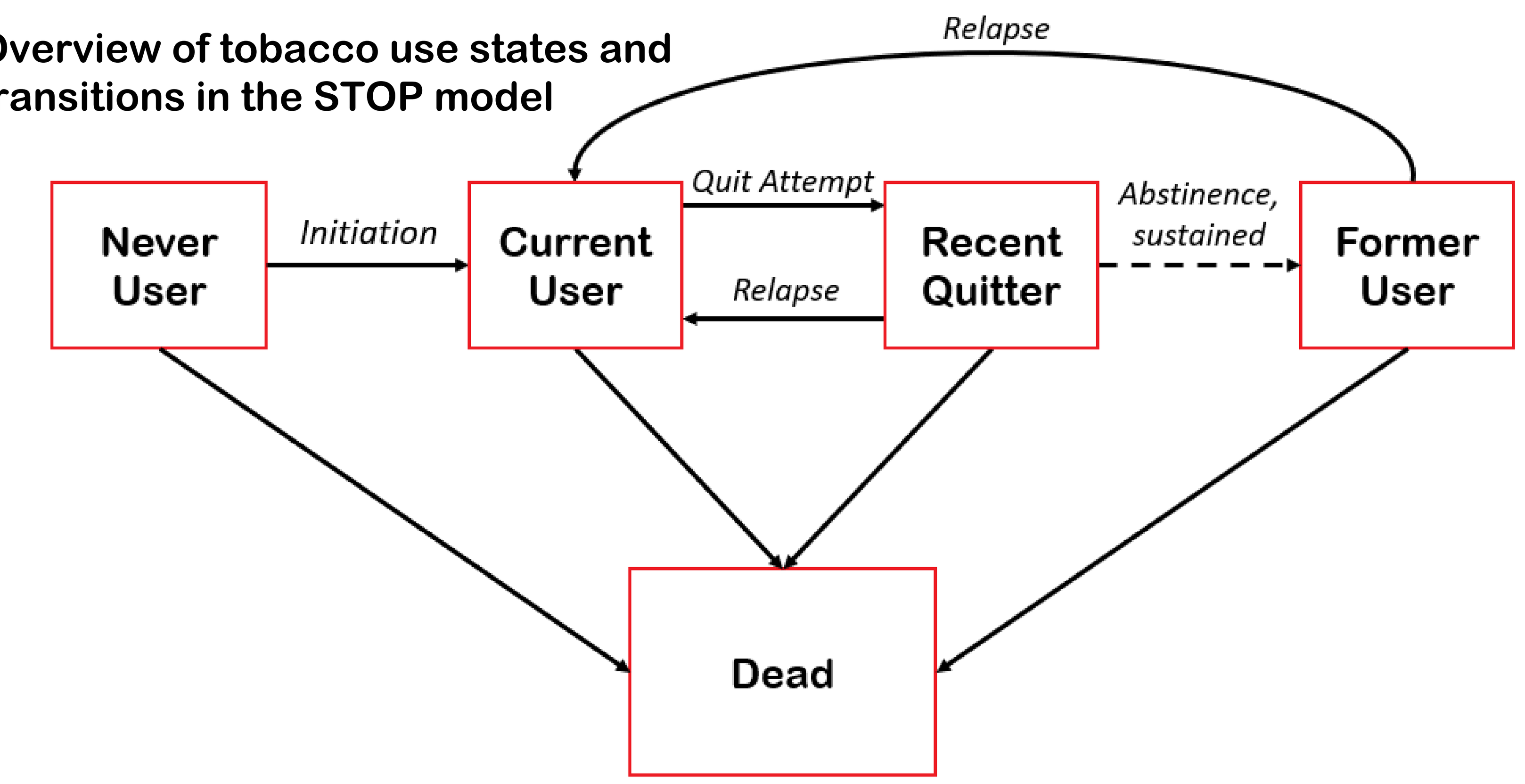
Data

ID	WAVE	SEX	AGE	RACE	EDUCATION	RELAPSE	WEIGHT	USE STATE
P00004	1	Male	18-24 years old	White Only	Some college	0	7073.69	NSFE
P00004	2	Male	18-24 years old	White Only	Some college	1	7073.69	NSCE
P00005	3	Female	45-64 years old	Other	Advanced Degree	0	3036.67	FSNE
. . .								
P23569	4	Female	12-14 years old	Black Only	High school or less	0	1522.36	NSNE
P23569	4.5	Female	15-17 years old	Black Only	High school or less	0	1522.36	NSNE
P23569	5	Female	15-17 years old	Black Only	High school or less	0	1522.36	NSCE

PATH is an on-going study aimed at providing an understanding of tobacco use patterns and related health effects in the U.S. Participants are followed up with over time (every one-two years) and their subsequent responses on various aspects of tobacco use are recorded. We synthesized data for **nine cigarette smoking and e-cigarette use states** (current/former/never use of each), and associated information on participant **sex, age, race, education,** and **history of relapse**. Responses were weighted with Wave 5 longitudinal weights to account for potential biases.

Methods

Overview of tobacco use states and transitions in the STOP model



Transition probabilities derived from the Markov multi-state model are given by:

$$P_{ij}(s,t) = P(X_t = j | X_s = i) \text{ for } i,j \in S, s \le t,$$

where S describes the state occupied at time t.

Covariate effects on transition rates were estimated. The estimate of the hazard, λ_{ij} is given by:

$$\lambda_{ij}(t) = \lim_{\Delta t \to 0} \frac{P(X_{t+\Delta t} = j | X_t = i)}{\Delta t},$$

where λ_{ij} is the instantaneous risk of moving from state i to state j. The transition intensity matrix gives the instantaneous rate of transition from one state to another :

$$Q(t) = \begin{pmatrix} -\lambda_{11} & \cdots & -\lambda_{19} \\ -\lambda_{21} & \cdots & -\lambda_{29} \\ \vdots & \ddots & \vdots \\ -\lambda_{91} & \cdots & -\lambda_{99} \end{pmatrix}$$

Transition probabilities are computed from these intensities as $\boldsymbol{P(t)} = \mathbf{exp}(\boldsymbol{Q(t)})$.

Given starting population estimates, these probabilities empower the STOP model to project prevalence of tobacco use into the future. These estimates are then compared against empirical PATH data for accuracy and goodness-of-fit.

STOP Model Projected Prevalences

	15-17	18-24	25-44	45-64	65+
Model Start	2.8%	19.0%	24.6%	20.0%	6.7%
Year 1	2.3%	16.8%	23.8%	20.7%	7.6%
Year 2	2.6%	15.8%	23.3%	20.8%	8.2%
Year 3	3.3%	15.4%	23.4%	20.9%	8.4%
Year 4	4.0%	14.5%	23.5%	20.9%	8.5%
Year 5	5.5%	14.2%	24.0%	21.0%	8.3%

Empirical PATH Prevalences

	15-17	18-24	25-44	45-64	65+
2015	2.8%	19.0%	24.7%	20.0%	7.7%
2016	2.4%	17.4%	24.1%	19.9%	8.3%
2017	1.7%	15.5%	24.5%	20.1%	8.3%
2018	1.6%	-	-	-	-
2019	1.5%	10.6%	23.3%	20.3%	8.2%

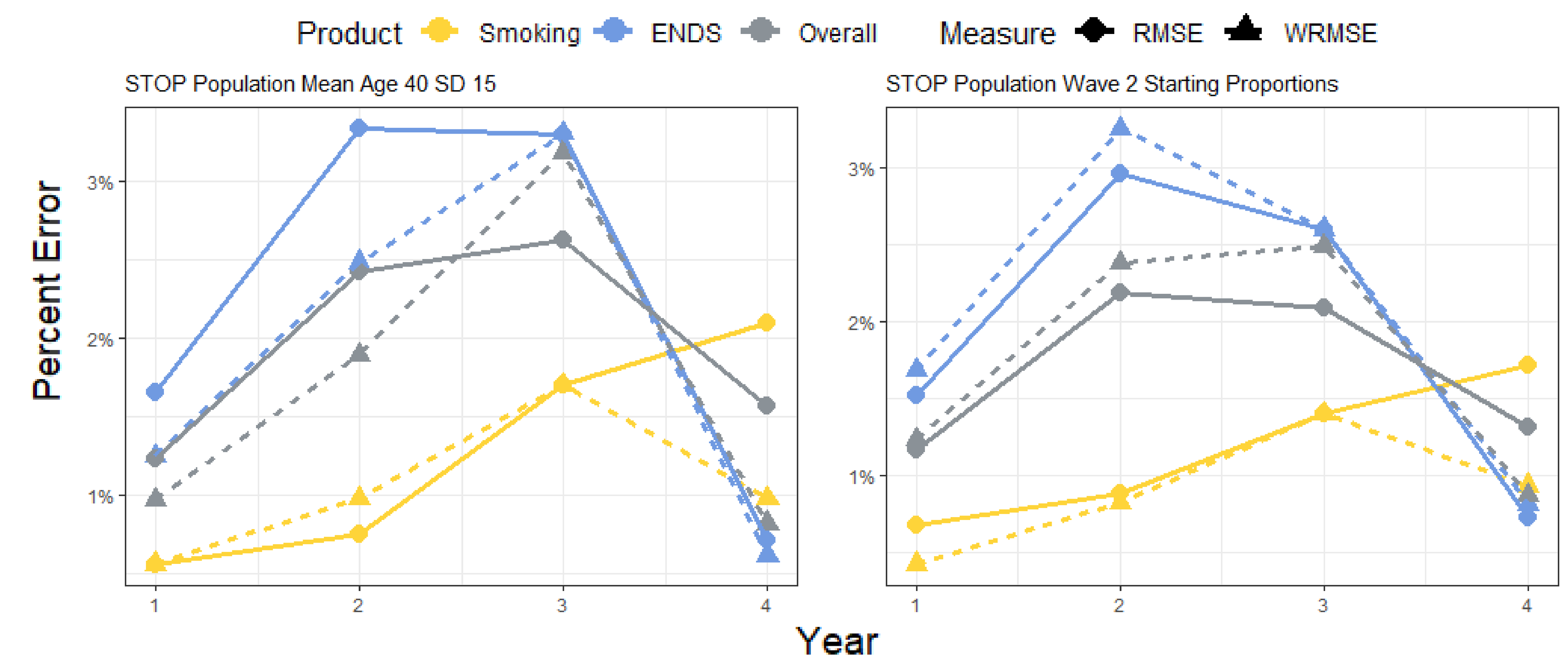
We fit mixed-effects logistic regression models to draw inference on the effects of our covariates on tobacco use:

$$logit(E[Current\ Tobacco\ Use_i \mid X_i]) = \beta_{0i} + \beta_1 Wave + \beta_2 Sex + \beta_3 Age + \beta_4 Race + \beta_5 Education + \beta_6 Num.Relapses,$$

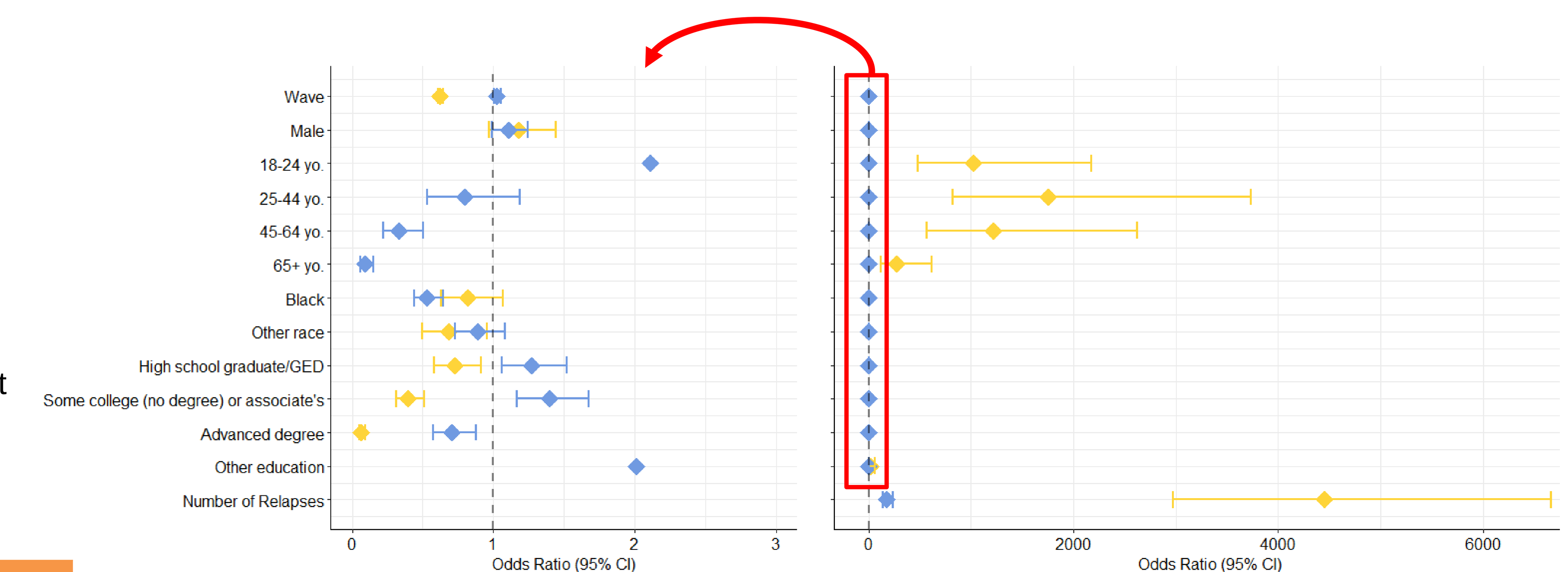
where β_{0i} is a random intercept for individual i in the study.

Current tobacco use is interchangeable with *current smoking* and *current ENDS* use of which we fit separate models for.

Results



Yearly STOP-projected prevalences of smoking and e-cigarette use were compared to empirical data from PATH in order to assess performance and goodness of fit. We tracked root-mean-squared-error (RMSE) and a weighted version of root-mean-squared error (WRMSE) based on the proportion of people represented within each group of the simulation, over time in model runs using a) an initial population drawn from a normal distribution with mean age 40 and standard deviation 15 (***left***) and b) an initial population informed by the proportion of people represented within each age group in Wave 2 of PATH (***right***). For both populations, STOP estimates of **cigarette use** more closely match empirical data compared to ENDS, with the exception of Year 4.



We extract odds ratios corresponding to our covariates of interest related to separate outcomes of current smoking (***yellow***) and current ENDS use (***blue***) with a reference group of *Female, 15-17 yo., White, in high school or lower, at a baseline time with 0 relapses observed during the duration of the study*. Notably, the effect of age for those ≥25y takes on a reverse effect depending on the tobacco product, further demonstrating the **popularity of e-cigarettes in younger populations** as opposed to older. ENDS are also highly prevalent among **high-school graduate/early college aged** individuals as opposed to cigarettes. Lastly, assuming a linear effect of time, while cigarettes have become *less* popular over time, ENDS have become slightly *more* popular over time, though this positive correlation is not significant.

Conclusion

- Youth tobacco use was generally more erratic than that of adults with these groups taking particular interest in **ENDS/e-cigarettes** over traditional cigarettes.
- Root-mean-squared error (RMSE) and weighted root-mean-squared-error (WRMSE) for STOP-projected versus PATH empirical prevalence for smoking and ENDS were **<4% for all four years** of simulation. We observed especially great performance on estimates for **Year 4/2019** (our primary target of interest due to the boom in JUUL use that occurred during this time).
- With confidence in our results, **tax policies** intended to influence tobacco use can be introduced into the STOP model, thereby providing decision makers with estimates of the potential impact of these policies.
- Mixed effects models revealed covariate relationships consistent with what was found in Markov multi-state modeling.