Research Article **Open Access**

Estimating Transitional Probabilities with Cross-Sectional Data to Assess Smoking Behavior Progression: A Validation Analysis

Xinguang Chen¹* and Feng Lin²

¹*Pediatrics Prevention Research Center/Department of Pediatrics, Wayne State University, Detroit, Michigan, USA*
²*Department of Electrical and Computer Engineering, Wayne State University, Detroit, Michigan, USA*

Abstract

Background and objective: New analytical tools are needed to advance tobacco research, tobacco control planning and tobacco use prevention practice. In this study, we validated a method to extract information from cross-sectional survey for quantifying population dynamics of adolescent smoking behavior progression.

Methods: With a 3-stage 7-path model, probabilities of smoking behavior progression were estimated employing the Probabilistic Discrete Event System (PDES) method and the cross-sectional data from 1997-2006 National Survey on Drug Use and Health (NSDUH). Validity of the PDES method was assessed using data from the National Longitudinal Survey of Youth 1997 and trends in smoking transition covering the period during which funding for tobacco control was cut substantively in 2003 in the United States.

Results: Probabilities for all seven smoking progression paths were successfully estimated with the PDES method and the NSDUH data. The absolute difference in the estimated probabilities between the two approaches varied from 0.002 to 0.076 (p>0.05 for all) and were highly correlated with each other (R^2=0.998, p<0.01). Changes in the estimated transitional probabilities across the 1997-2006 reflected the 2003 funding cut for tobacco control.

Conclusions: The PDES method has validity in quantifying population dynamics of smoking behavior progression with cross-sectional survey data. The estimated transitional probabilities add new evidence supporting more advanced tobacco research, tobacco control planning and tobacco use prevention practice. This method can be easily extended to study other health risk behaviors.

Keywords: PDES method; Validation; Adolescent smoking; Cross-sectional data

Introduction

Need for longitudinal data on smoking behavior progression

Exposure to tobacco is associated with 5.4 million deaths per year worldwide [1] and among them, 440,000 deaths are in the United States [2]. Despite much progress in tobacco use prevention across the globe [3], further advancement in tobacco control requires data beyond the static measures of prevalence rates to cover steps of smoking behavior progression. Smoking behavior is rather complex and it involves a series of sophisticated neurobiological, psychosocial and behavioral processes [4-7]. Researchers who have investigated the population dynamics of smoking behavior proposed a number of models, including but are not limited to, Flay's five-stage model (preparatory, trying, experimental, regular use and addicted/dependent use) [8], which was adopted by the U.S. Department of Health Services in the 1994 Surgeon General's Report [9]; Prochaska's Theory of Stages of Change and Transtheoretical Model [10,11]; and Mayhew's six-stage conceptual framework that integrates both Flay's and Prochaska's models [12]. Although these models are promising, establishment of such models requires panel data collected through longitudinal designs.

Collecting longitudinal data involves repeatedly following-up of individual participants over time. This is technically demanding and practically expensive because of increased efforts to plan and implement such projects and increased burdens on the participants (being repeatedly contacted) and the related entities (e.g., families, schools, etc) [13,14]. Consequently, longitudinal surveys are relatively scant, and such data particularly lack in the developing and transitional counties with more than 80% of the smokers in the world but limited resources for tobacco research and tobacco control planning and practice [15]. Nevertheless, more longitudinal information is needed

(1) to better understand the population dynamics of smoking behavior, (2) to locate strategically sensitive steps (e.g., smoking initiation, quitting, relapse, etc) along the smoking behavior progression for prevention intervention, and (3) to evaluate a prevention program on various progression steps for improvement.

Challenges to extracting longitudinal information from cross-sectional data

Cross-sectional data are widely available from a number of sources, including the Global Youth Tobacco Survey, the National Survey on Drug Use and Health (formerly known as National Household Survey on Drug Abuse), the Youth Risk Behavior Surveillance, and Monitor the Future Studies. Although no individual participants are followed in a cross-sectional survey, conceptually data from such surveys may contain longitudinal information. For example, a cross-sectional survey of a sample of participants (*a*) to (*a+n*) years old can be considered equivalent to a longitudinal survey that follows a sample of participants, all (*a*) years of age for *n* years. Likewise, such a cross-sectional survey is conceptually also equivalent to a longitudinal survey with two waves of

*Corresponding author: Xinguang Chen, MD, PhD, Professor, Pediatric Prevention Research Center/Department of Pediatrics, Wayne State University School of Medicine, 4707 St. Antoine Street, Hutzel W534, Detroit, Michigan 48201, USA, Tel: (313) 745-0564; E-mail: jimchen@med.wayne.edu

data collection in two consecutive years for participants aged (*a* to *n*-1) at the first and aged (*a*+1 to *n*) at the subsequent year.

Conceiving cross-sectional surveys in a longitudinal perspective creates a basis which supports the development of methods capable of extracting longitudinal information from such data [16]. For example, if from a cross-sectional survey conducted in one year, we estimated that the number of neversmokers for participants aged 12 and 13 are 400 and 300 respectively, then the probability for a participant aged 12 to remain as a never-smoke during an one-year period would be approximately 0.75 (=300/400). However, computing transitional probabilities by this approach requires two additional assumptions: (1) Changes in the number of people by age between two consecutive years are negligible (stable population assumption as in demographic studies for life expectancy estimation) [16] and (2) changes in smoking behavior for individuals of the same age between two consecutive years are also negligible compared to changes in smoking behavior across ages in a year.

PDES Method as an alternative

To formalize the approach, we use the theory of Probabilistic Discrete-Event Systems that offers a method by which longitudinal information can be derived from cross-sectional data. The PDES method is an established analytical technique for modeling and control in industry to describe assembly lines and other systems [17-20]. According to the PDES method, a cross-sectional survey is like a "snapshot" of the status of a system and the dynamics of a system can be described with such snapshot data. In a previous study, we have mathematically established the PDES method that can minimize the impact of population variations in births, deaths and substantial changes in smoking prevalence on modeling by the use of state probabilities [21]. In this study, we reported our empirical work to validate the PDES method.

The PDES method differs from a few methods that have been attempted by others [22,23]. In one study that focused on prediction of smoking using a state-transition model, the transition rates of starting, quitting and relapsing were estimated using cross-sectional data and a restricted quadratic multinomial and quadratic logistic regression spline [22]. However, this method did not consider changes in population and smoking behavior. Another study employed a heterogeneous Markov model to estimate entry and exit transition probabilities, but for this method to work, data from at least two consecutive cross-sectional surveys are needed [23]. In addition, none of these methods have been validated.

Methods

Data for PDES modeling

Ten-year cross-sectional data were derived from the National Survey on Drug Use and Health (NSDUH) collected during the 1997-2006. The NSDUH is an on-going effort sponsored by the Substance Abuse and Mental Health Services Administration and carried out by the Research Triangle Institute, Cary, North Carolina through contracted projects. The multi-stage random cluster sampling scheme was used in the NSDUH to select participants that represent all civilian and non-institutionalized population 12 years of age and older in the United States. Participants 12-17 years old were included for this analysis. After a screening test, the trained data collectors were sent to the sampled households to administrate the survey using the Computer-Assisted Personal Interviewing (CAPI) technique. The 1997 NSDUH data were used for comparison with the probabilities estimated with longitudinal

data (see the section on longitudinal data later in this paper) and the 10-year NSDUH data were used to show time trends in smoking behavior progression. Although changes were made to the NSDUH in 1999 (increased sample size) and 2002 (introduced incentives to adolescents), previous studies showed limited impact of these changes on the overall trend of cigarette smoking [24].

PDES method and smoking progression

According to the PDES method, modeling an assembly system with a number of connected workstations can be achieved through cross-sectional assessment (snapshot) [22,23]. When the system is in running parts for a product (e.g., a car) are continuously put into the system and processed through various workstations to produce the needed product. Using PDES, such an assembly system G can be described as:
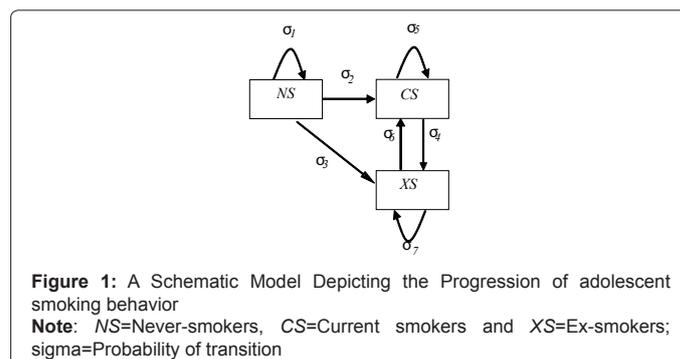
$$G=(Q, \Sigma, \delta, q_0) \tag{1}$$

where $Q=\{q_0,q_1,\ldots,q_n\}$ is the state set of the system such as idle, working or breakdown; $\Sigma=\{\sigma_1,\sigma_2,\ldots,\sigma_m\}$ is the event set representing transitions from one state to another; $\delta:Q\times\Sigma\rightarrow Q$ is the transition function describing what event can occur at which state and the resulting states; and $q_0 \in Q$ is the initial state. For example, δ (idle, startup)=working means that at "idle" stage, the event "startup" will bring the system to "working". To simplify the notations, we also use q_i to denote the probability of the system being at state q_i and use σ_i to denote the probability of σ_i occurring.

To describe the system, multiple cross-sectional measures (snapshots) of the system status Q will be obtained: $Q(t)$, $Q(t+T)$, $Q(t+2T)\ldots$, (T=sampling interval). The PDES method assumes that the system status at time $t+T$ depends on its status at time t and the transitions occurred during the sampling interval [$t, t+T$]. Obviously, when the system is running in a stable status, one snapshot would contain adequate information to describe the system.

Following the assembly principle, a 3-stage model (Figure 1) was proposed to validate the PDES method in analyzing smoking behavior. In the model, *NS*=never-smokers, participants who have never smoked by the time of survey; *CS*=current smokers, participants who smoked currently; and *XS*=ex-smokers, participants who ever smoked and did not smoke now, therefore $Q=\{NS, CS, XS\}$. The arrowed lines in the figure indicate the seven transition paths or events, $\Sigma=\{\sigma_1,\sigma_2\ldots,\sigma_7\}$. When individual children in a population grow up, they will all pass through the PDES system to become different types of smokers. Likewise, data from a cross-sectional survey provides a snapshot $Q=\{NS, CS, XS\}$, from which transitional probabilities $\sigma_1, \sigma_2, \sigma_3 \ldots \sigma_7$ can be estimated (Figure 1).

PDES model is an extension of discrete event systems model. When considering probabilities of states and events, PDES models are similar
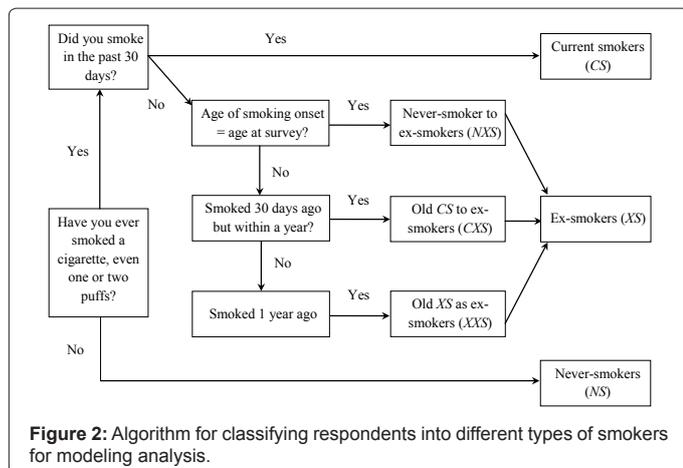

Figure 1: A Schematic Model Depicting the Progression of adolescent smoking behavior
Note: *NS*=Never-smokers, *CS*=Current smokers and *XS*=Ex-smokers; sigma=Probability of transition

Figure 2: Algorithm for classifying respondents into different types of smokers for modeling analysis.

to MC models. However, the PDES models also consider properties such as controllability, observability, detectability, and diagnosability that are not considered in MC models.

Categorization of participants into different types of smokers

Figure 2 illustrates the algorithm used in this study to classify participants into various types of smokers for analysis. To define $Q=\{NS, CS, XS\}$, data from the two questions were used. (1) "Have you ever smoked a cigarette, even one or two puffs?" (yes/no) and (2) "How long has it been since you last smoked a cigarette?" (within 30 days, 30 days to one year, one year ago and within three years, and three years ago). Participants were classified as NS if they responded negatively to question (1); participants will be coded as CS if they smoked within the past 30 days based on their responses to questions (1) and (2); participants were classified as XS if they responded positively to question (1) and last smoking was 30 days ago.

To solve for the PDES model, three additional types of smokers were needed: (1) NXS -- NS progressed to CS and further to XS within the past year; (2) CXS -- CS a year ago (old CS) progressed to XS in the past year; and (3) XXS -- XS a year ago (old XS) remained as XS in the past year (see Figure 2 for details). To specify these three types of smokers, data from one more question "How old (age in years) were you the first time you smoked a cigarette, even one or two puffs?" was added.

Estimation of state probability and transitional probability

The state probability for each of the six smoking types described above was computed as the proportion (%) of a type relative to the total sample. The state probabilities were computed by age to obtain $NS(a)$, $NXS(a)$, $CS(a)$, $XS(a)$ for a=12, 13…17 and were used as data to solve for $\Sigma=\{\sigma_1(a),\sigma_2(a),…,\sigma_7(a)\}$. As an innovation of the PDES method, utilization of the state probabilities minimizes the impact of sudden changes in population (births, deaths) and/or smoking behavior on transitional probability estimation [21]. To account for the complex sampling designs used in NSDUH and NLSY97, the Proc Survey Means was used to compute state probabilities.

To obtain transitional probabilities, the estimated state probabilities by single year of age were converted to state probability at the beginning of an age by taking an average of two probabilities at the consecutive age groups. For example, $[NS(12)+NS(13)]/2 \approx NS(13)$ is the state probability at the beginning of age 13. With the converted state probabilities, the following matrix equation was used to estimate all transitional probabilities $\Sigma=\{\sigma_1(a),\sigma_2(a),…,\sigma_7(a)\}$:

$$\begin{bmatrix} \sigma_1(a) \\ \sigma_2(a) \\ \sigma_3(a) \\ \sigma_4(a) \\ \sigma_5(a) \\ \sigma_6(a) \\ \sigma_7(a) \end{bmatrix} = \begin{bmatrix} 0 & -NS(a) & -NS(a) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & CS(a) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & XS(a) \\ 0 & 0 & NS(a) & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} NS(a+1)-NS(a) \\ CXS(a+1) \\ XXS(a+1) \\ NXS(a+1) \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Longitudinal data for validation analysis

Longitudinal data for validating the PDES model were derived from the National Longitudinal Survey of Youth 1997 (NLSY97). The NSLY97 is sponsored by the Bureau of Labor Statistics, U.S. Department of Labor and carried out by the Center for Human Resource Research, Ohio State University. Participants were selected using the multi-stage stratified random sampling method. Youths 12 to 16 years old by the end of 1996 were eligible to participate and they became 12 to 18 (mean=14, SD=1.3) in 1997 after completing the baseline survey (n=8,984), a response rate was 72%. Data were collected by trained researchers at home using the CAPI technique. In addition to cigarette smoking, survey dates were collected for assessing the duration between surveys.

Estimation of transitional probabilities with longitudinal data

We directly estimated the same seven transitional probabilities $\Sigma=\{\sigma_1(a),\sigma_2(a),…,\sigma_7(a)\}$ by age as with the PDES method with the NLSY97 data. In computing these transitional probabilities, we first defined the three types of smokers NS, CS, and XS at the baseline in 1997 and the follow-up in 1998 respectively. This was conducted following the definitions described for the PDES method above and data derived from the following three questions: (1) "Have you ever smoked a cigarette?" (2) During the past 30 days, on how many of the days did you smoke a cigarette?" (3) "Have you smoked a cigarette since the last interview on [date of last interview]?"

With the number of NS, CS, and XS by age in 1997 and 1998, transitional probability from one type of smokers to another during the one-year period was estimated as the ratio of the two in 1997 and 1998. Since the time interval between the two surveys for individual participants was not equal but varied from 6-23 months, the method of person-years at-risk was used for probability estimation [13,14].

Data processing and statistical computing

Both the NSDUH data and NLSY97 data were acquired through the Inter-University Consortium for political and social research. Data were re-coded after a thorough review of all the related technical documents from the data provider. The commercial software SAS 9.2 (SAS Institute Inc, Cary, NC) was used for data processing and general statistical analysis. Matlab was used to solve the matrix PDES equation.

Results

Sample characteristics and state probability

Table 1 summarizes basic demographic characteristics of the study samples. Data in the upper panel of the table indicates that the participants of the NSDUH varied from 8,731 in 1997 to 20,838 in 2006 with the response rates varying from a minimum of 61.4% in 1999 to a maximum of 78.3% in 1997. These participants, 50% male and more than 50% white, were 12-17 years old with a mean age of 14.8-15.0 (SD=1.9 to 2.0). Data in lower panel of the table indicates that among 8,984 participants of the NLSY97 aged 14.4 (SD=1.5) at baseline, 8386 (93%) participated in follow-up survey when they were on average 16.0 years old.

Advances in Markov Chain Monte Carlo
Methods and Survival Analysis

	Total	Mean age (SD)	% Male	% White	% Black	%Hispanic	% Response Rate
NSDUH							
1997	8731	14.9 (1.9)	49.4	50.6	17.2	26.6	78.3
1998	7880	15.0 (2.0)	49.9	45.0	20.9	27.6	77.0
1999	21197	14.9 (1.9)	50.4	67.2	13.4	13.5	61.4
2000	21982	14.9 (1.9)	50.7	66.8	13.5	13.9	68.6
2001	19854	14.9 (2.0)	50.4	66.9	13.3	13.2	67.4
2002	20106	14.8 (2.0)	50.9	66.5	13.5	13.8	71.3
2003	20834	14.9 (2.0)	51.5	63.0	14.2	14.7	70.7
2004	20980	14.9 (2.0)	50.9	63.9	13.4	14.4	70.1
2005	21241	14.9 (1.9)	50.5	62.1	13.7	16.0	69.8
2006	20838	15.0 (1.9)	51.2	60.8	14.1	16.6	68.4
NLSY97							
1997	8984	14.4(1.5)	51.2	49.6	26.2	10.9	91.6
1998	8386	16.0(1.4)	51.1	49.3	26.5	10.9	93.3*

Note: NSDUH: the National Survey on Drug Use and Health, NLSY97: the National Longitudinal Survey of Youth 1997. The cross-sectional 1997 NSDUH data were used for estimating probabilities of smoking behavior transitions with the Probabilistic Discrete Event System (PDES) method; the longitudinal NLSY97 data with baseline data collected in 1997 and followed up data collected in 1998 were used to estimate transitional probabilities for the same smoking behavior transitions; and a comparison of the two sets of transitions provided evidence assessing the validity of the PDES method. The 10-year NSDUH data were used to estimate smoking behavior transitions for US youth 12-17 years of age by the validated PDES method. *: One-year follow-up rate of the NLSY97.

Table 1: Characteristics of the Study Samples from the 1997-2006 National Survey on Drug Use and Health and the 1997 National Longitudinal Survey of Youth 1997.

Age (year)	Sigma1	Sigma 2	Sigma 3	Sigma 4	Sigma 5	Sigma 6	Sigma 7
PDES method and the cross-sectional 1997 NSDUH data[a]							
13	0.880	0.070	0.050	0.193	0.807	0.476	0.526
14	0.859	0.074	0.068	0.174	0.826	0.437	0.563
15	0.836	0.089	0.076	0.187	0.814	0.400	0.600
16	0.863	0.066	0.071	0.196	0.804	0.430	0.570
17	0.891	0.045	0.046	0.148	0.853	0.442	0.558
Directly estimated with the 1997 National Longitudinal Survey data[b]							
13	0.885	0.067	0.048	0.159	0.841	0.425	0.575
14	0.889	0.059	0.052	0.115	0.885	0.409	0.591
15	0.889	0.065	0.046	0.068	0.932	0.387	0.613
16	0.873	0.083	0.044	0.088	0.912	0.426	0.574
17	0.884	0.073	0.043	0.080	0.920	0.423	0.577
Differences between the two estimations (13-17 years old)							
Mean	0.002	0.020	0.076	0.076	0.020	0.020	0.016
SD	0.023	0.012	0.038	0.038	0.016	0.016	0.024
P (t test)	>0.05	>0.05	>0.05	>0.05	>0.05	>0.05	>0.05

Note: [a]Transitional probabilities were estimated with one-year cross-sectional data from the 1997 National Survey on Drug Use and Health. [b]Transitional probabilities were estimated with data collected among youth who participated the baseline and the follow-up survey from the National Longitudinal Survey of Youth 1997. See Figure 1 and the corresponding text for definitions of the seven transitional probabilities. *:This sample contained only participants who provided completed data at both baseline survey in 1997 and follow-up survey in 1998.

Table 2: Transitional Probabilities Estimated with PDES Method/1997 NSDUH data (N=8,731) and the Conventional Method and the NLSY97 Data (n=7,286*).

Table 2 contains two sets of probabilities for the seven smoking progression steps, one set was estimated with the cross-sectional 1997 NSDUH data and the other set was estimated with the longitudinal NLSY97 data. According to the results from NSDUH data, an adolescent aged 13 who have never smoked in 1997 has 88.0% (σ_1=0.880) chance to remain as a never-smoker in a year; an adolescent in the same age who was smoking has 19.3% (σ_4=0.193)

chance to quit; an ex-smoker in the same age has 47.6% (σ_6=0.476) chance to relapse and re-smoke.

A visual assessment of all results in (Table 2) indicates that the two sets of probabilities estimated by different methods and data are close to each other. For example, the estimated σ_1=0.880 for adolescents of 13 years of age when the PDES method was used, and the same σ_1=0.885 when the longitudinal method was used. The estimated σ_6 across ages from 13 to 17 from the PDES method varied from 0.400 to 0.476, close to those from the longitudinal method that varied from 0.387 to 0.426. Data in the bottom row of (Table 2) indicate the mean differences between the two sets of estimates varied from 0.002 (SD=0.023) to 0.076 (SD=0.038), and none of them were statistically significant (t-test, p>0.05 for all). When the two sets of the estimated probabilities were cross-plotted, they were distributed closely around the diagonal (data are not shown) with a very high correlation (R^2= 0.988, p<0.001).

Figure 3 depicts the time trends for the five key transitional probabilities (σ_2, transition from never-smokers to current smokers; σ_3, transition from never-smokers to ex-smokers; σ_4, quitting or transition from current smokers to ex-smokers; σ_6, relapsing or transition from ex-smokers to current smokers; and $\sigma_2+\sigma_3$, rate of smoking initiation). While changes over time in σ_2, σ_3, as well as ($\sigma_2+\sigma_3$) were relatively small, the probability for quitting (σ_4) showed an increasing trend from 1997 to 2002 before it declined suggesting more smokers quitting before 2002 while fewer smokers quitting since 2003. A trend opposite to σ_4 was observed for σ_6, the probability of relapsing with fewer ex-smokers relapsing before 2003 and more ex-smokers relapsing since 2003.

Discussion

PDES method is valid for examination of adolescent smoking

Findings of this study suggest certain validity of the newly established PDES method [21] in estimating transitional probabilities with cross-sectional survey data. Although only one-wave survey data was used, the estimated probabilities from the PDES method are very close to those computed from longitudinal data. In addition, changes in the estimated transitional probabilities from cross-sectional data were closely associated with the event of funding cut for tobacco control in the United States, including funding cut from the Master Settlement Agreement [25]. In addition to validating the method, more obvious changes in quitting and relapsing than in other progression steps imply higher sensitivity of these two progression steps than other steps in
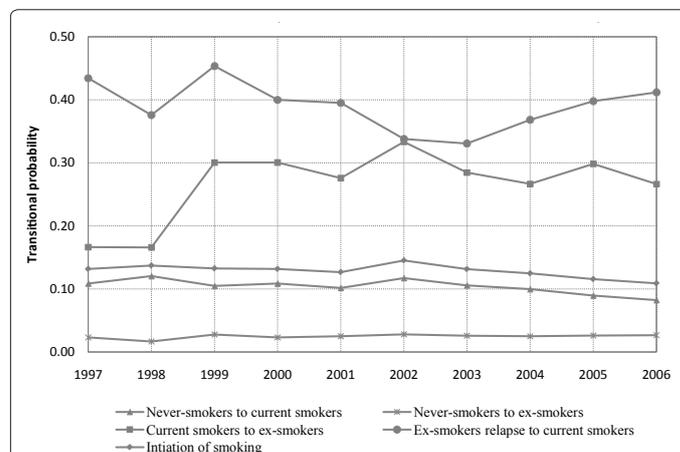
Figure 3: Probabilities of smoking behavior transitions among US adolescents, 1997-2006, (PDES method and 1997-2006 NSDUH data).

response to substantive funding cut for tobacco use prevention. This useful evidence would not be revealed without the PDES method. Although PDES method is not intended to replace the longitudinal method, it adds an alternative approach for tobacco research, tobacco control planning and tobacco use prevention practice. This method will be of particular significance to countries and places where tobacco use is highly prevalent while resources are limited to collect longitudinal data [26,27].

Technical consideration in the application of the PDES method

Cross-sectional survey data are widely available to assess tobacco use at the state, national, and even global levels [15,28-32]. With the PDES method, transitional probabilities can be estimated with data from one single cross-sectional survey to assess various steps of smoking behavior progression over time, by single year of age, and stratified by gender, race/ethnicity to describe population dynamics of tobacco use behavior. Transitional probabilities can also be estimated across subgroups of significant predictor variables such as education, peer influences, school performance, parental monitoring, and receptivity to pro-tobacco media to assess factors associated with smoking behavior progression.

One advantage of the PDES method is that it can be used to assess effect from tobacco control at the macro level. When multi-year cross-sectional data are available, like the NSDUH used in this study, time trends of the estimated transitional probabilities with such data can be related to various tobacco control activities (e.g., tobacco taxation, legal restrictions, school-based programs, or tobacco cessation) to assess effects from such effort. As we have showed in this study that reductions in quitting and increase in relapsing in the US were associated with sustained effort for tobacco control up to 2002 before a sudden and substantial funding cut at the national and state level in 2003. Such effect could not be revealed without application of the PDES method [24,33]. In addition to historical analysis, a contrast of transitional probabilities between exposed and non-exposed youth will provide data for assessing tobacco control programs. For example, data from the NSDUH on exposure to several types of educational programs [34,35] can be used for such purpose.

When transitional probabilities are compared over time or across interventions, information can also be derived to assess: (1) tobacco control effort on specific steps of smoking behavior progression (e.g., from never-smokers to smoker and further to ex-smokers); (2) the effect in a specific smoking progression step (e.g. increasing quitter or reducing experimenters) in reducing the total number of smokers; (3) the progression steps that are sensitive to change; and (4) the amount of changes needed in a transition step to achieve a pre-determined tobacco control objective. This type of information is useful for tobacco control planning and program strategy optimization [15,36,37].

Despite the strengths, caution is suggested when the PDES method is used when sudden and substantial changes in population size smoking behavior in the survey year. In addition, we recommend using single-year age-group for analysis so that the stability assumption of the PDES method will not be violated.

Acknowledgements

References

1. World Health Organization (2008) WHO Report on the Global Tobacco Epidemic 2008: The MPOWER Package. Geneva.

2. U.S. Department of Health and Human Services (2004) The Health Consequences of Smoking - A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

3. Warren CW, Jones NR, Eriksen MP, Asma S; Global Tobacco Surveillance System (GTSS) collaborative group. (2006) Patterns of global tobacco use in young people and implications for future chronic disease burden in adults. Lancet 367: 749-753.

4. Chassin L, Presson CC, Pitts SC, Sherman SJ (2000) The natural history of cigarette smoking from adolescence to adulthood in a midwestern community sample: multiple trajectories and their psychosocial correlates. Health Psychol 19: 223-231.

5. Chen X, Stacy A, Zheng H, Shan J, Spruijt-Metz D, et al. (2003) Sensations from initial exposure to nicotine predicting adolescent smoking in China: a potential measure of vulnerability to nicotine. Nicotine Tob Res 5: 455-463.

6. Kandel DB, Kiros GE, Schaffran C, Hu MC (2004) Racial/ethnic differences in cigarette smoking initiation and progression to daily smoking: a multilevel analysis. Am J Public Health 94: 128-135.

7. Robinson ML, Berlin I, Moolchan ET (2004) Tobacco smoking trajectory and associated ethnic differences among adolescent smokers seeking cessation treatment. J Adolesc Health 35: 217-224.

8. Flay BR (1993) Youth tobacco use: Risks, patterns, and control. In Nicotine Addiction: principles and management, Oxford University Press, New York.

9. U.S. Department of Health and Human Services (1994) Preventing Tobacco Use among Young People - A Report of the Surgeon General: Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

10. Pallonen UE, Prochaska JO, Velicer WF, Prokhorov AV, Smith NF (1998) Stages of acquisition and cessation for adolescent smoking: An empirical integration. Addict Behav 23: 303-324.

11. Prochaska JO, DiClemente CC (1983) Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 51: 390-395.

12. Mayhew KP, Flay BR, Mott JA (2000) Stages in the development of adolescent smoking. Drug Alcohol Depend 59: S61-S81.

13. Gordis L (1996) Epidemiology. Philadelphia/London/Tronto: W.B. Saunders Company.

14. Rothman KJ, Greenland S, Lash TL (2008) Modern Epidemiology. (3rdedn), Wolters Kluer/Lippincott Willimans & Wilkins,Philadelpha/Baltimore/New York/London.

15. Warren CW (2008) The Global Youth Tobacco Survey (GYTS): linking data to the implementation of the WHO Framework Convention on Tobacco Control. BMC Public Health 8: S1.

16. Keyfitz N, Caswell H (2005) Applied Mathematical Demography. (3rdedn), Springer Science, New York.

17. Kim DS, Kulkarni DM, Lin F (2002) An upper bound for carriers in a three-workstation closed serial production system operating under production blocking. IEEE Transactions on Automatic Control 47: 1134-1138.

18. Li Y, Lin F, Lin ZH (1999) Supervisory control of probabilistic discrete-event systems with recovery. IEEE Transitions on Automatic Control 44: 1971-1975.

19. Lin F (1994) A linguistic approach to stochastic comparison of queuing-systems. European Journal of Operational Research 77: 27-39.

20. Shu S, Lin F, Ying H, Chen X (2008) State estimation and detectability of probabilistic discrete event systems. Automatica (Oxf) 44: 3054-3060.

21. Lin F, Chen X (2010) Estimation of transitional probabilities of discrete event

systems from cross-sectional survey and its application in tobacco control. Inf Sci (Ny) 180: 432-440.

22. Capannesi M, Boshuizen HC, Willemsen MC, van Houwelingen HC (2009) How to obtain long term projections for smoking behaviour: A case study in the Dutch population. Comput Math Methods Med 10: 155-164.

23. Pelzer B (2002) Bayesian estimation of transition probabilities from repeated cross sections. Statistica Neerlandica 56: 23-33.

24. Chen X, Lin F, Stanton B, Zhang X (2011) APC modeling of smoking prevalence among US adolescents and young adults. Am J Health Behav 35: 416-427.

25. Niederdeppe J, Farrelly MC, Hersey JC, Davis KC (2008) Consequences of dramatic reductions in state tobacco control funds: Florida, 1998-2000. Tob Control 17: 205-210.

26. Roemer R, Taylor A, Lariviere J (2005) Origins of the WHO Framework Convention on Tobacco Control. Am J Public Health 95: 936-938.

27. Taylor AL, Bettcher DW (2000) WHO Framework Convention on Tobacco Control: a global "good" for public health. Bull World Health Organ 78: 920-929.

28. Braverman MT, Svendsen T, Lund KE, Aaro LE (2001) Tobacco use by early adolescents in Norway. Eur J Public Health 11: 218-224.

29. Chen X, Li G, Unger JB, Liu X, Johnson CA (2003) Secular trends in adolescent never smoking from 1990 to 1999 in California: an age-period-cohort analysis. Am J Public Health 93: 2099-2104.

30. Moolchan ET, Fagan P, Fernander AF, Velicer WF, Hayward MD, et al. (2007) Addressing tobacco-related health disparities. Addiction 102: 30-42.

31. Rudatsikira E, Muula AS, Siziya S (2009) Current cigarette smoking among in-school American youth: results from the 2004 National Youth Tobacco Survey. Int J Equity Health 8: 10.

32. White VM, Hayman J, Hill DJ (2008) Can population-based tobacco-control policies change smoking behaviors of adolescents from all socio-economic groups? Findings from Australia: 1987-2005. Cancer Causes Control 19: 631-640.

33. Chen X, Lin F, Jiang Y (2010) PDES Modeling of Smoking Behavior Transition and National Trends in Cigarette Smoking among US Adolescents-A Novel Approach to Smoking Behavior Analysis. Presented at the the 138th Annual Meeting of American Public Health Association Denver, CO.

34. SAMHSA (2006) National Survey on Drug Use & Health Bethesda, Maryland.

35. Carpenter CS (2007) Workplace drug testing and worker drug use. Health Serv Res 42: 795-810.

36. Mamudu HM, Hammond R, Glantz SA (2008) Project Cerberus: tobacco industry strategy to create an alternative to the Framework Convention on Tobacco Control. Am J Public Health 98: 1630-1642.

37. Murnaghan DA, Leatherdale ST, Sihvonen M, Kekki P (2008) A multilevel analysis examining the association between school-based smoking policies, prevention programs and youth smoking behavior: evaluating a provincial tobacco control strategy. Health Educ Res 23: 1016-1028.