



## Review

## Monitoring the tobacco use epidemic III The host: data sources and methodological challenges

Cristine D. Delnevo<sup>a,\*</sup>, Ursula E. Bauer<sup>b</sup><sup>a</sup> UMDNJ-School of Public Health, 335 George Street, Suite 2100, New Brunswick, NJ 08901, USA<sup>b</sup> New York State Department of Health, New Brunswick, NJ 08901, USA

## ARTICLE INFO

Available online 25 September 2008

## Keywords:

Tobacco control  
Surveillance  
Evaluation  
Survey

## ABSTRACT

**Objective.** This Host paper (III of V) reviews key surveillance and evaluation systems that monitor the characteristics, attitudes and behaviors of tobacco users that are crucial for tobacco control efforts.

**Methods.** We summarize and expand on the recommendations from the Host Working Group of the National Tobacco Monitoring, Research and Evaluation Workshop. We also discuss research challenges and make additional recommendations for improving tobacco control surveillance and evaluation.

**Results.** We reviewed 10 major US surveys that collect data on tobacco use. A great deal of data is collected but gaps exist. Data collection on cigars, smokeless tobacco, brand, menthols, and PREPs is sparse and infrequent. Also, a number of factors, including, but not limited to, changes in US population composition, declines in survey response rates, and increases in cell phone use present research challenges that may impact the ongoing utility of these systems.

**Conclusions.** Although the field of tobacco control research is an advanced area of public health, improvements in data systems are necessary to accurately evaluate progress and continue tobacco control gains. A coordinated surveillance and evaluation network would increase efficiency and improve the overall utility, quality and timeliness of the current data systems.

© 2008 Elsevier Inc. All rights reserved.

## Contents

Introduction . . . . .	S17
Methods: Inventory of constructs and existing activities . . . . .	S17
National Health Interview Survey (NHIS) . . . . .	S17
Behavioral Risk Factor Surveillance System (BRFSS) . . . . .	S17
Tobacco Use Supplement to the Current Population Survey (TUS CPS) . . . . .	S18
National Health and Nutrition Examination Survey (NHANES) . . . . .	S19
National Survey on Drug Use and Health (NSDUH) . . . . .	S19
Health Information National Trends Survey (HINTS) . . . . .	S19
Adult Tobacco Survey (ATS) . . . . .	S19
The Youth Risk Behavior Surveillance System (YRBSS) . . . . .	S19
Youth Tobacco Survey (YTS) . . . . .	S19
Monitoring the Future (MTF) . . . . .	S19
Discussion: Methodological challenges and opportunities . . . . .	S20
Rapid response surveillance . . . . .	S20
Surveying special populations . . . . .	S20
Wireless substitution . . . . .	S20
Declining response rates . . . . .	S20
Setting effects . . . . .	S21
Parental consent . . . . .	S21
Standardization . . . . .	S21
Analytic complexity . . . . .	S21
Single vs. multi-topics surveys . . . . .	S21

\* Corresponding author. Fax: +1 732 235 9777.

E-mail address: [delnevo@umdnj.edu](mailto:delnevo@umdnj.edu) (C.D. Delnevo).

Repeated cross sectional vs. cohort/panel surveys. . . . .	S21
Multi-level data and modeling. . . . .	S22
Conclusion . . . . .	S22
Conflict of interest statement . . . . .	S22
Acknowledgments . . . . .	S22
References . . . . .	S22

## Introduction

Understanding, documenting and quantifying the characteristics of the tobacco user, or potential user, has been key to tobacco control efforts. The tobacco user is considered the “host” in the traditional public health epidemiology model of disease. Although the field of tobacco control is an advanced area of public health, improvements in data systems are necessary to evaluate progress and continue tobacco control gains (Giovino et al., 2009), particularly given the highly-adaptive nature of the vector (Cruz, 2009), addictive properties of the agent (Stellman and Djordjevic, 2009), and variability in environmental influences (Farrelly, 2009). After reviewing the current state of host-related surveillance and evaluation, and identifying key gaps and challenges, the Host Working Group at the 2002 National Tobacco Monitoring, Research, and Evaluation Workshop made the four priority recommendations shown in Table 1. This paper reviews and updates the above recommendations and discusses challenges and opportunities with respect to the research, design, measurement and analysis of tobacco-related host characteristics.

## Methods: Inventory of constructs and existing activities

A variety of monitoring, research and evaluation systems exist to collect information on host characteristics at the national, state, and local levels. The surveys and systems assess tobacco-related behaviors, attitudes and perceptions, and often collect information on other lifestyle and health-related behaviors. While each survey serves a particular purpose, addressing a specific surveillance or research need, no individual survey is without gaps and weaknesses. In 2001, CDC summarized basic information on such data systems and produced a brief compilation of data sources useful to tobacco control programs (Yee and Schooley, 2001). The inventory reported here builds on this work by providing methodological details on specific surveys and updates the earlier publication by including the most current surveys in use. Also, the inventory consists only of items from ongoing systems implemented at the national level or at the state level with coordination by a federal agency. Many states (e.g., California, Massachusetts), organizations (e.g., the American Legacy Foundation), and research programs have developed data systems to meet specific information needs. Also, international efforts, such as the International Tobacco Control Policy Evaluation Project (ITC) bear mentioning with respect to its prospective cohort study design, scope (i.e., multiple countries), and focus on the Framework Convention on

Tobacco Control's (FCTC) policies (Fong et al., 2006). Although the inventory does not include items from these systems, it should be noted that such systems both inform and are informed by national surveys. A brief description of each survey's design and their host content areas follows; host survey measures are summarized in Table 2. While the focus of this paper is the host, it is important to note that most of the surveys and systems described also contain measures on the environment, as discussed in a companion paper by Farrelly (2009).

### National Health Interview Survey (NHIS)

Conducted by the CDC's National Center for Health Statistics, the NHIS has been a primary source of annual health data on the US population since the 1950 s; tobacco measures have been incorporated on 29 NHIS administrations since 1965. Details on the NHIS design are found elsewhere (Centers for Disease Control and Prevention, 2008a). In brief, the NHIS uses a multistage area probability sample of the civilian non-institutionalized population of the US. Sample weights adjust for the complex sample design. Since 1997, NHIS has used computer-assisted personal interviewing (CAPI) to administer the Adult Core questionnaire to approximately 31,000 individuals annually (ages > 18 years) in their homes, with overall response rates ranging from 72% to 75%. The 2006 NHIS Adult Core questionnaire assessed cigarette smoking, age of initiation, consumption, quit attempts, and abstinence. Additionally, NHIS addresses health insurance coverage, other health behaviors (e.g., alcohol use, physical activity), pregnancy status, and morbidity. NHIS is one of the few surveys presented in the host inventory that addresses mental health. Mental health measures have been included via the K6 scale, a brief scale for nonspecific psychological distress in NHIS since 1997 (Kessler et al., 2002).

Detailed questions on tobacco use are included in a Cancer Control Supplement to the NHIS, which has occurred about every 5 years since 1987 (1987, 1992, 2000, 2005). The 2005 Cancer Control Supplement addressed the use of other tobacco products (e.g., cigars, pipes, smokeless tobacco), menthol cigarettes, readiness to quit, cessation methods, and exposure to secondhand smoke in the home. Perceived risks of secondhand smoke and opinions about tobacco control policies were collected in the 2000 supplement, but not in 2005.

### Behavioral Risk Factor Surveillance System (BRFSS)

The BRFSS, established by the CDC in 1984, collects data on health behaviors from a representative sample of civilian non-institutionalized adults in each of the 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. Details on BRFSS's design are found elsewhere (Centers for Disease Control and Prevention, 2008b). In brief, BRFSS is a state based random-digit-dial (RDD) computer-assisted-telephone-interview (CATI) survey of the adult (ages > 18 years) population. Disproportionate stratified samples are commonly used by states and the use of sample weights is required. In recent years, the total sample size has increased substantially. In 2007, a total of 430,912 adults participated and state sample sizes ranged from 2,552 to 39,549; the median response rate was 50.8% and ranged from 26.9% to 65.4% (Center for Disease Control and Prevention,

**Table 1**

Priority recommendations made by the host working group at the 2002 national tobacco monitoring, research, and evaluation workshop

- 1) Complete an inventory of existing surveys and questions and identify the extent to which host tobacco issues are addressed.
- 2) Identify methodologies for rapid response surveillance and issues and characteristics that can be addressed by a rapid response mechanism;
- 3) Develop appropriate instruments and sampling methodologies to assess tobacco use attitudes and behaviors in special populations.
- 4) Establish a coordinated national tobacco surveillance system that can serve as a clearinghouse for government and nonprofit agencies and organizations involved in monitoring, research and evaluation of tobacco use and prevention programs.

**Table 2**  
Inventory of tobacco host related measures across surveys

Topic	2007 NHIS	2005 NHIS CC Sup.	2007 BRFSS	2005 ATS		2006 CPS-TUS	2007 NSDUH	NHANES	2005 HINTS	2007 MTF	2006 NYTS	2007 YRBS
				C	M							
<i>Cigarette smoking behaviors</i>												
Ever puffed/tried					X		X		X	X	X	X
100 Lifetime	X		x	x		x	x	x	x		x	
Now smoke every day or some days	X		x	x		x	x	x	x		x	x
Past 30 days	X			x		x	x			x	x	x
Age of first use	X				x	x	x	x		x	x	x
Number smoked per day	X			x		x	x	x	x	x	x	x
Quit attempts/history	X		x	x		x	x	x	x	x	x	x
Duration of abstinence/former smokers	X			x		x	x	x	x			
Cotinine								x				
Brand(s)					x		x	x		x	x	
UPC Code					x			x				
Regular/Light/Ultra-light					x		x	x			x	
Menthol		X			x	x	x	x			x	
Discount					x		x	x				
<i>Use of other tobacco products</i>												
Smokeless tobacco		X			x	x	x	x		x	x	x
Cigars		X			x	x	x	x			x	x
Pipes		X			x	x	x	x			x	
Bidis or Kreteks (clove cigarettes)		X			x		x	x		x	x	
Roll-your-own							x					
<i>Youth issues</i>												
Susceptibility							x			x	x	
School performance							x			x		
Body weight concerns										x		x
<i>Other</i>												
Home smoking bans										x		
Readiness/motivation to quit		X			x	x			x	x	x	
Addiction/dependence indicators					x	x	x	x		x	x	
Self-efficacy about quitting						x					x	
Methods used to quit		X			x	x		x			x	
Perceived health risks/benefits of smoking					x				x	x	x	
Perceived risks of ETS					x						x	
Perceptions about tobacco products						x			x		x	
Use of PREPS						x	x		x	x		
Opinions about tobacco control policies						x						
Opinions about tobacco companies						x						
<i>Health behaviors/status/consequences</i>												
Alcohol and drug use and abuse	X		x				x	x		x		x
Other health behaviors	X		x					x		x		x
Pregnancy status	X		x				x	x				
mental Mental health indicators	X		x				x	x	x			
Morbidity	X		x					x				
Health insurance coverage	x		x					x				

2008b). The questionnaire addresses a variety of health behaviors and issues and is comprised of core questions and optional modules that states utilize based on need. Historically, BRFSS included a notable amount of host content, such as the number of cigarettes smoked each day by current smokers, use of other tobacco products, physician assistance, and secondhand smoke policies, in either the core or optional modules. However, more recent surveys have contained only a few tobacco host items. Indeed, in 2007, core questions consisted of only two measures: cigarette use and quit attempt in past 12 months and states were not offered an optional tobacco module. Given the burden of tobacco-caused disease in the United States, the erosion of tobacco content in the BRFSS is problematic.

#### *Tobacco Use Supplement to the Current Population Survey (TUS CPS)*

The CPS is a labor force survey conducted monthly by the U.S. Census for the Bureau of Labor Statistics. The Tobacco Use Supplement (TUS), sponsored by the National Cancer Institute (NCI) since 1992 and co-sponsored by NCI and CDC since 2001, is fielded approximately

every 3 years (National Cancer Institute, 2008a). The CPS uses an area probability sampling design to select a stratified probability sample of clusters of households. Approximately 56,000 households are surveyed in a given month using CATI and CAPI methods. State estimates may be generated from the national TUS by combining multiple months of data, with total sample sizes of respondents (> 15 years old) in excess of 240,000; state samples range from 2500 (Mississippi) to 17000 (California). Individual level response rates for the TUS-CPS questionnaire are approximately 81–85% (including self- and proxy respondents and 65–72% for self-response only) among those households completing the basic CPS household survey (response rates range from 93 – 97%). The Special Cessation Supplement in 2003 and the 2006–07 TUS-CPS questionnaire contain numerous items relevant to the tobacco host, including cigarette use, age of initiation, number of cigarettes smoked, quit attempts, use of other tobacco products, use of menthol cigarettes, cessation and addiction indicators, utilization of cessation services, advice from health professionals to quit smoking, attitudes toward smoking in various public places, and use of “potential reduced exposure products” (PREPs).

### National Health and Nutrition Examination Survey (NHANES)

The NHANES, conducted by the CDC, is a key data source on the health and nutritional status of the civilian non-institutionalized U.S. population. Historically, surveys were conducted on a periodic basis from 1960 to 1994; in 1999 NHANES became a continuous survey. The detailed methodologies of the NHANES are available elsewhere (Centers for Disease Control and Prevention, 2008c). In brief, the sample is selected through a complex, multistage area probability design. The survey includes an interview conducted in the household and an examination at a mobile examination center, which collects laboratory data. Sample sizes for analyses will vary with the population of interest. For example only those ages 16+ complete the smoking and tobacco use questionnaire; approximately 5,000 individuals participated in this component of the 2003–2004 NHANES. A considerable number of tobacco host measures are contained in NHANES, including cigarette use, age of initiation, number of cigarettes smoked, use of other tobacco products (i.e., cigars, smokeless tobacco, and pipes), detailed brand information (e.g., name, light cigarettes, menthol cigarettes), addiction indicators, alcohol and drug use/abuse, pregnancy status, and mental health indicators. NHANES' unique contribution is the collection of biological samples which allows calculation of cotinine level for biochemical validation of smoking status (Caraballo et al., 2004, Caraballo et al., 2001), or exposure to secondhand smoke (Pirkle et al., 2006) as well as assessment of smokers' cotinine levels (O'Connor et al., 2006).

### National Survey on Drug Use and Health (NSDUH)

Formerly known as the National Household Survey on Drug Abuse, the NSDUH has collected data from residents of households, non-institutional group quarters, and civilians living on military bases since 1971 (Substance Abuse and Mental Health Services Administration, 2006). In 2005, the NSDUH underwent a redesign resulting in a coordinated sample plan that will remain in use until 2009; the design utilizes independent stratified, multistage, area probability samples within each of the 50 states and the District of Columbia. Each year approximately 67,500 individuals aged 12 and older are surveyed via a computer-assisted interview. The sample is equally distributed among three age groups: 12 to 17 years, 18 to 25 years, and 26 years or older. In 2006, NSDUH reported a screening response rate of 91% and an interview response rate of 74%. The 2006 questionnaire contains many host tobacco measures, including cigarette use, age of initiation, other tobacco products, menthol, light and discount cigarettes, addiction indicators, alcohol and drug use/abuse, pregnancy status, and mental health indicators measured via the K6 scale (Kessler et al., 2002). It is the only survey described herein that collects brand name information not only for cigarettes, but for cigars and smokeless tobacco. The NSDUH has undergone several major methodological changes (e.g., name change, sampling, method of data collection, incentives), which may impact survey estimates (Gfroerer et al., 2002; Kennet and Gfroerer, 2005) thus trends only since 2002 should be reported (Kennet and Gfroerer, 2005).

### Health Information National Trends Survey (HINTS)

Recently, NCI initiated HINTS to assess the prevalence of cancer-relevant knowledge, attitudes, and behaviors. HINTS is intended to be an ongoing, cross-sectional survey of the U.S. civilian, non-institutionalized, adult (18+) population. Details on the design of HINTS are found elsewhere (National Cancer Institute, 2008b). In brief, participants are sampled via a RDD design and minorities are over-sampled; sample weights are utilized for analyses. HINTS I occurred in 2003; a total of 6,369 adults participated and the overall response rate was 33%. HINTS II occurred in 2005; 5,586 adults participated in HINTS II and the overall response rate was 20.8%. Survey measures relative to

the host include cigarette use, number of cigarettes smoked, quit attempts, attitudes towards and use of PREPs, and mental health indicators measured via the K6 scale (Kessler et al., 2002). Thus far, HINTS has received limited attention from tobacco control researchers, who should be encouraged to consider this new data source.

### Adult Tobacco Survey (ATS)

The surveys discussed thus far collect some information on host tobacco factors, but do not cover the topic in depth. In 2002, CDC, in conjunction with state partners, developed the Adult Tobacco Survey to obtain comprehensive information for state tobacco control programs. The ATS design, largely based on the BRFSS, is a state based RDD CATI survey of the adult (ages 18+) non-institutionalized population with telephones. Disproportionate stratified samples are often used and sample weights are required to generate representative estimates. The ATS questionnaire is comprised of core questions and optional modules that states utilize based on data needs; state-added questions are also incorporated. Since 2003, 19 states have conducted a CDC-supported ATS, for a total of 32 unique administrations. Sample sizes have ranged from 1,301 to 9,179, and response rates ranged from 15.7% to 50.3%. However, state ATS data is proprietary, no system exists to disseminate the data, and survey content and methods can vary widely from state to state.

### The Youth Risk Behavior Surveillance System (YRBSS)

The YRBSS includes national, state and local surveys. Coordinated by the CDC, the YRBSS has been monitoring health risk behaviors among high school students, biennially, for over a decade. The details of the YRBSS design are described elsewhere (Brener et al., 2004). In brief, the national survey (NYRBS) utilizes a three-stage cluster sample design to obtain a nationally representative sample of students in grades 9 through 12, whereas state and local YRBSSs employ a two-stage design to produce representative samples of students in grades 9–12 in their jurisdictions. Sample weights are utilized. The YRBS is self-administered in a school setting and addresses six priority health risk behaviors, including tobacco use. The 2007 NYRBS survey contains 87 items, including the following host tobacco measures: experimentation and current use of cigarettes, cigars and smokeless tobacco products, number of cigarettes smoked, quit attempts, and use on school grounds. The YRBSS provides comparable tobacco host data on youth at the local, state, and national levels.

### Youth Tobacco Survey (YTS)

CDC developed the YTS in 1998 to provide data to support the design, implementation, and evaluation of state tobacco control programs. The YTS includes state, national, and international school-based surveys of middle and high school students. The National and State YTS and YRBS use identical sampling methodologies and the same wording for shared survey items; they are administered in alternating years. The YTS supplements the YRBS by providing more comprehensive tobacco-related information on both middle school and high school students. The 2006 NYTS instrument contained 81 items. The YTS has considerable coverage of host measures, including use of tobacco products, cigarette brand, menthol cigarettes, consumption patterns, cessation and addiction indicators, susceptibility to smoking, and perceptions (see Table 2).

### Monitoring the Future (MTF)

Begun in 1975, MTF is a research program of the Institute for Social Research at the University of Michigan (Johnston et al., 2005). In brief, MTF employs a multistage sampling design to obtain nationally representative samples of secondary school students (i.e., 8th-, 10th-,

and 12th-grade students) from the 48 contiguous states. Data have been collected annually from 12th grade students since 1975 and from 8th- and 10th-grade students since 1991. At present, approximately 50,000 students in about 420 public and private secondary schools are surveyed annually. The school participation rate ranges from 66% to 85% and the student response rates range from 79% to 91%. Sample weights are utilized to generate representative estimates. Since 1976, a representative sample of the 12th grade students has participated in a longitudinal panel. Host tobacco content in recent years includes use of cigarettes and smokeless tobacco, quit attempts, susceptibility, perceptions regarding tobacco use, purchasing patterns, school performance, and use of alcohol and drugs. MTF has remained methodologically consistent over time, making it an excellent data source to examine trends over the past 30 years.

While there is a great deal of host tobacco data collected in the US, gaps exist for host measures. Few surveys collect data on brand, use of menthols, lights and ultra-lights, and PREPs. With an evolving tobacco market, product related data are sorely needed. Also, coverage of other tobacco products, most notably cigars and snuff, two products with increasing consumption (US Department of Agriculture, 2006, Delnevo, 2006) is sparse and data collection infrequent.

### Discussion: Methodological challenges and opportunities

The surveys described herein have historically worked well to provide valuable tobacco host data. However, a few key methodological challenges and opportunities have the potential to influence the ongoing utility of these systems and are briefly described below. Moreover, it is important to stress that responding to these challenges requires balancing the utility of surveillance systems with the need for consistent methods and measures over time.

#### Rapid response surveillance

The ever-evolving, and sometimes rapidly changing, landscape of the tobacco policies, and tobacco industry innovation and promotions requires timely assessments of factors known to impact behavior. Thus, the Host working group sought to potential methodologies for such rapid response studies. One emerging survey methodology worth attention is internet panel surveys, which offer a promise of rapid turnaround and low cost. However, more research is needed with respect to issues pertaining to sampling frames and generalizability.

In addition, some current surveys can theoretically be adapted to provide information rapidly. For example, in the weeks following the attacks of September 11th, three states added a terrorism module to BRFSS within 4 weeks of the attacks and collected data from over 3500 participants in 12 weeks (Centers for Disease Control and Prevention, 2002). Another recent benefit of BRFSS is that weighted datasets are available to states on a quarterly basis, enabling timely feedback, analysis, and dissemination. However, modifying BRFSS to add items of emerging interest on short notice can be challenging and has not yet been attempted as a coordinated effort nationally.

#### Surveying special populations

The US is increasingly diverse with respect to race/ethnicity and almost 1 out of 5 people speak a language other than English at home (U.S. Bureau of the Census, 2003). The tobacco industry has a history of marketing its products disproportionately to members of minority communities and marginalized groups (Apollonio and Malone, 2005; Smith and Malone, 2003; U.S. Department of Health and Human Services, 1998). Accordingly, a long term recommendation identified by the Host working group was to develop appropriate instruments and sampling methodologies for special populations. These populations include, but are not limited to minorities, non-English speaking,

the Gay, Bisexual, Lesbian, or Transgender (GLBT) communities, the homeless, prison populations, and individual with disabilities. The long term nature and complexity of this recommendation requires continuous discussion and feedback and precludes providing specific details here. However, several points related to survey development and sampling bear mention and should be considered in future surveys. First, in recent years federal agencies, such as the US Census Bureau and NCI have recognized that collecting quality data from populations where English is a second language requires more than survey translation, such as back-translation, and cognitive testing (Pan and de la Puente, 2005; Willis et al., 2005). Second, sampling methodologies must address the fact that certain groups may be geographically clustered or not accessible via traditional sampling frames, such as those living in group quarters, including both institutional (e.g., correctional institutions, nursing homes, mental hospitals) and non institutional group quarters (dorms, military quarters and homeless shelters).

Citizens of low and middle income countries also bear mention given that the burden of tobacco-caused disease falls disproportionately in such countries (Jha and Chaloupka, 2000). Research and surveillance initiatives such as the Global Tobacco Research Network (GTRN), co-sponsored by NCI, the Global Tobacco Surveillance System (i.e., Global Youth Tobacco Survey, Global School Personnel Survey, and Global Health Professions Student Survey) supported by CDC and WHO, and the recent Bloomberg Initiative to develop and implement a Global Adult Tobacco Survey in developing countries are vital to inform programming to curb the global tobacco epidemic.

#### Wireless substitution

The percent of adults who lived in cell phone or wireless only households more than quadrupled between early 2003 (2.9%) and late 2007 (14.5%) and existing indicators suggest wireless substitution will continue to grow (Blumberg and Luke, 2008). Most random digit dial (RDD) telephone survey do not include cell-phone numbers, thus excluding adults in wireless households. Specific subpopulations more likely to live in a wireless household include young adults, lower income, renters, males, minorities, and smokers (Blumberg et al., 2007, 2008). Even after adjusting for demographic variations, the exclusion of the cell-phone only population biases cigarette smoking estimates by 1 percentage point or greater (Blumberg et al., 2008). In due course, the effect of cell phone-only households could mask or exaggerate real variations in smoking prevalence.

Recent research using BRFSS data provides evidence that wireless substitution may be biasing and underestimating health indicators (i.e., current smoking and drinking), for young adults (Delnevo et al., 2008) Young adults have been historically challenging to sample and survey given their high mobility rate and the exclusion of college dorms in most RDD surveys. Paradoxically, while wireless substitution may create challenge for traditional RDDs, the trend may actually prove to be beneficial with respect to sampling young adults. The high rate of cell phone ownership among this cohort, combined with number portability, may minimize or even potentially eliminate prior sampling challenges for this population.

#### Declining response rates

Survey response rates have fallen across all modes in recent years, most dramatically among RDD surveys (Hox and de Leeuw, 1994, Steeh et al., 2001). Indeed, over a decade, the median response rate for BRFSS has declined from 69.9% in 1994 to 50.5% in 2007 (Centers for Disease Control and Prevention, 2008b). Because response rates are an indicator of survey data quality, such declines raise concern about whether estimates are biased. A study by Biener and colleagues examined the impact of declining response rates on smoking estimates in California and Massachusetts and concluded that despite

declines in response rates, prevalence estimates were not compromised (Biener et al., 2004). While these findings are encouraging, efforts should continue to maximize response rates to facilitate generalizability, and further research is called for to examine potential bias in estimates due to low response rates or survey drop out. Differential nonresponse by specific population groups (e.g., smokers, minorities) has the potential to introduce bias in smoking estimates, but this has not been well studied or documented.

#### *Setting effects*

Collecting data about youth tobacco use is crucial to tobacco control efforts, but existing surveys often do not produce similar results. The setting for data collection (i.e., home vs. school) is one factor which may affect the accuracy of youth risk behavior estimates. The context of setting is important as it relates to the extent of privacy allowed during data collection. Also, while school-based surveys can be anonymous, home surveys cannot. The lack of anonymity and/or privacy may affect disclosure of risky behavior. Household surveys tend to produce lower tobacco use estimates than school based surveys, although these differences are not always significant (Gfroerer et al., 1997, Kann et al., 2002). Recent work using automated telephone interviewing (i.e. respondent responds to a question by pressing numbers on telephone key pad) indicates that youth were more likely to report smoking when provided with the privacy afforded by this emerging technology (Currivan et al., 2004). A recent study, utilizing the YRBS, compared setting as well as mode of data collection (computer vs. paper) and found that for those behaviors with a significant setting effect, the school-based survey yielded higher estimates than home (Brener et al., 2006). Setting effects were strongest for illegal or socially stigmatized behaviors; the effect on tobacco estimates was only significant for current smokeless tobacco use.

#### *Parental consent*

School based surveys require some form of parental consent. The type of procedure used (e.g., active or passive) may impact participation rates, which, in turn, can affect the generalizability of findings. Active consent procedures often yield lower response rates than passive or parental notification procedures (Baker et al., 2001, Dent et al., 1993). When active consent is used the highest risk youth can be less likely to participate, resulting in participation bias (Dent et al., 1993, Eaton et al., 2004). Some studies have found lower cigarette smoking estimates when active consent procedures are used (Anderman et al., 1995, Pokorny et al., 2001). which may be attributable to low response rates. However, an analysis of NYRBS found that parental permission procedures do not affect prevalence estimates when high response rates are obtained (Eaton et al., 2004).

#### *Standardization*

One use of the surveys described herein is to provide information relevant to the evaluation of tobacco control programs. Measuring and quantifying intervention effects at the state level is easier when national estimates or those from other states can be used as a comparison to demonstrate unique program effects. Such comparisons benefit from standardization. Cigarette smoking measures in adult surveys, and to some extent in youth surveys, have been fairly well standardized, allowing for examination of trends overtime within surveys as well as comparisons between surveys. It should be noted that this applies only to English language surveys; basic cigarette smoking questions vary in their wording for the Spanish language versions of BRFSS, TUS-CPS and NHIS. Moreover, standardization for other host measures, such as use other tobacco products, has lagged (National Cancer Institute, 1998). Over time, new resources have

emerged to facilitate comparability across surveys. These resources draw from the existing systems, including one-time surveys, and provide a foundation for the development of new surveys and systems. An example of such a resource is the CDC's "Question Inventory on Tobacco."

#### *Analytic complexity*

Many of the surveys discussed in the inventory have complex sampling design (e.g., stratification, clusters, multi-stage samples), resulting in unequal probabilities of selection. Such designs complicate the analysis since observations are neither independent nor selected with equal probability. Historically, the use of standard statistical analysis packages (e.g., SAS, SPSS) to analyze complex surveys resulted in underestimated standard errors and so the analysis has traditionally been the purview of specialized software, such as SUDAAN or WesVar. However, recent versions of SAS, SPSS and STATA include procedures to address complex samples. These packages produce results similar to SUDAAN (Siller and Tompkin, 2005), providing multiple choices for tobacco researchers. Access to appropriate data analysis tools increases the likelihood that information from monitoring, research and evaluation systems will be used locally to inform tobacco control programs.

#### *Single vs. multi-topics surveys*

A recent concern in tobacco monitoring is that tobacco specific surveys (i.e., ATS, YTS) tend to produce different, and usually lower, estimates of tobacco use behaviors than multi-topic surveys (e.g., BRFSS, YRBS) (Cowling et al., 2003, Kalsbeek et al., 2003 Ramsey et al., 2004, RTI International, 2006). In some cases, reported differences in smoking prevalence have been substantial, as much as 5 percentage points (Kalsbeek et al., 2003; Ramsey et al., 2004). Research in this area is limited and several reasons for the differences have been posited, including but not limited to tobacco-specific survey introduction, and question context and ordering. There is evidence to support that the tobacco-specific survey introduction of the ATS may cue potential participants into the topic of the survey and prompt current smokers to either deny their tobacco use (Cowling et al., 2003) or refuse participation in the survey (RTI International, 2006), thereby lowering the resulting prevalence estimates of tobacco use. Kalsbeek and colleagues sought to examine difference between NYTS and NYRBS and considered the impact of question context on youth smoking prevalence. In this case, tobacco questions on the NYRBS were preceded by more sensitive topics (e.g., violence, sexual behavior, and illegal drug use). The study found higher prevalence rates in the NYRBS compared to the NYTS, but the only statistically significant finding, as it related to question context was for ever cigarette smoking (Kalsbeek et al., 2003.). Given the remaining uncertainty on this issues, more methodological work is needed. Researchers should use caution when modifying existing surveillance systems or shifting to new systems, as these changes can produce changes in tobacco use prevalence estimates that are not reflective of actual changes in the population.

#### *Repeated cross sectional vs. cohort/panel surveys*

All of the surveys/systems discussed herein utilize a repeated cross-sectional design. The Monitoring the Future study is the one exception, in that it also includes a longitudinal panel. Prospective cohort or longitudinal panel designs offer a key methodological advantage over repeated cross sectional surveys with respect to causality. Such studies can better assess tobacco use trajectories (e.g., initiation, cessation, relapse) and determine the effect of tobacco control policies and programs as well as "vector" activities on tobacco use. The following longitudinal studies/surveys contain at least some

tobacco host measures: the International Tobacco Control Policy Evaluation Project (Fong et al., 2006), Panel Study on Income Dynamics (University of Michigan, 2008), National Longitudinal Study of Adolescent Health (University of North Carolina-Carolina Population Center, 2008), and the National Youth Cessation Survey (Klein et al. in press). However, it is important to point out that such surveys are not without their own challenges, including high costs and perhaps most importantly issues of attrition and retention. Indeed, retention in the ITC's US sample has been more challenging than their UK, Canadian, and Australian counterparts (Fong et al., 2006).

### Multi-level data and modeling

Tobacco use behavior is multi factorial and while the data systems aimed towards monitoring the host are individual level, other systems addressing influencing factors, such as the environment (e.g., smoke free policies) are not, and are often aggregated at geographic levels. Consideration must be given as to the extent to which these can be linked. For example, BRFSS and TUS CPS data can be linked with corresponding state, and sometimes county, level environmental data. However, indicators of geographic location are often suppressed in the public access files for many of the other data systems, such as NSDUH and NHIS, which introduces barriers to data analysis and dissemination. Also, it must be recognized that multi level data increases the complexity of analyses and appropriate statistical methodologies and models should be considered for such hierarchical data.

### Conclusion

A systematic understanding of the tobacco host remains unavailable due to gaps in survey content, limited monitoring efforts for specific populations, lack of coordination among researchers, and limited dissemination. As such, the Host working group calls for the establishment of a coordinated national tobacco surveillance system that can serve as a clearinghouse and address the challenges described herein. Such a network requires, at a minimum, cataloging and sharing data collection instruments and methodologies; establishing familiarity with, and enhancing existing systems; developing rapid response systems to meet emerging needs; the establishment of national data repository; and the resources to establish the network. A priority of the proposed network must be to more fully utilize collected data to inform public health practice. In sum, a comprehensive coordinate national system of ongoing monitoring, research and evaluation of tobacco-related host characteristics should facilitate: monitoring of tobacco control program progress, identification of gaps and weaknesses in tobacco control efforts, targeting and tailoring of interventions to specific populations, and countering tobacco industry practices.

### Conflict of interest statement

The authors declare that there are no conflicts of interest.

### Acknowledgments

The authors thank Lois Biener, Anne Hartman, and Gary Giovino for their comments on previous drafts; and the participants in the Host Working Group, Ursula Bauer, Ph.D., M.P.H.; Vilma Cokkinides, Ph.D., M.S.P.H., M.S.; Suzanne Colby, Ph.D.; Pebbles Fagan, Ph.D.; Elizabeth Gilpin, M.S. (Chair); Corinne Husten, M.D., M.P.H.; Deirdre Lawrence, Ph.D., M.P.H.; and Saul Shiffman, Ph.D.

### References

Anderman, C., Cheadle, A., Curry, S., Diehr, P., Shultz, L., Wagner, E., 1995. Selection bias related to parental consent in school-based survey research. *Eval. Rev.* 19, 663–674.

- Apollonio, D.E., Malone, R.E., 2005. Marketing to the marginalised: tobacco industry targeting of the homeless and mentally ill. *Tob. Control* 14, 409–415.
- Baker, J.R., Yardley, J.K., McCaul, K., 2001. Characteristics of responding-, nonresponding- and refusing-parents in an adolescent lifestyle choice study. *Eval. Rev.* 25, 605–618.
- Biener, L., Garrett, C.A., Gilpin, E.A., Roman, A.M., Curren, D.B., 2004. Consequences of declining survey response rates for smoking prevalence estimates. *Am. J. Prev. Med.* 27, 254–257.
- Blumberg, S.J., Luke, J.V., 2008. Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2007. Available at: <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless200805.pdf>. Accessed May 30, 2008.
- Blumberg, S., Luke, J.V., Cynamon, M.L., 2007. Telephone coverage and health survey estimates: Is concern about wireless substitution warranted? *Am. J. Public Health* 96, 926–931.
- Blumberg, S., Luke, J.V., Cynamon, M.L., Frankel, M.R., 2008. Recent trends in household telephone coverage in the United States. In: Lepkowski, T., Brick, ea (Eds.), *Advances in Telephone Survey Methodology*. John Wiley and Sons.
- Brener, N.D., Kann, L., Kinchen, S.A., et al., 2004. Methodology of the youth risk behavior surveillance system. *Morbidity and mortality weekly report. Recomm. Rep.* 53, 1–13.
- Brener, N.D., D, K.E., Kann, L., Grunbaum, J.A., Gross, L.A., Kyle, T.M., 2006. The association of survey setting and mode with self-reported health risk behaviors among high school students. *Public Opin. Q.* 70, 354–374.
- Caraballo, R.S., Giovino, G.A., Pechacek, T.F., Mowery, P.D., 2001. Factors associated with discrepancies between self-reports on cigarette smoking and measured serum cotinine levels among persons aged 17 years or older: Third National Health and Nutrition Examination Survey, 1988–1994. *Am. J. Epidemiol.* 153, 807–814.
- Caraballo, R.S., Giovino, G.A., Pechacek, T.F., 2004. Self-reported cigarette smoking vs. serum cotinine among U.S. adolescents. *Nicotine Tob. Res.* 6, 19–25.
- Centers for Disease Control and Prevention, 2002. Psychological and emotional effects of the September 11 attacks on the World Trade Center-Connecticut, New Jersey, and New York, 2001. *MMWR- Morbidity and Mortality Weekly Report* 51 (35), 784–786.
- Centers for Disease Control and Prevention, 2008a. National Health Interview Survey. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention URL <http://www.cdc.gov/nchs/nhis.html>. Accessed January 23, 2008.
- Centers for Disease Control and Prevention, 2008b. Behavioral Risk Factor Surveillance System. Atlanta Georgia: U.S. Department of Health and Human Services Available at: <http://www.cdc.gov/brfss>. Accessed March 18, 2008.
- Centers for Disease Control and Prevention (2008c). National Health and Nutrition Examination Survey. Available at <http://www.cdc.gov/nchs/nhanes.html>. Accessed January 23, 2008
- Cowling, D.W., Johnson, T.P., Holbrook, B.C., Warnecke, R.B., Tang, H., 2003. Improving the self reporting of tobacco use: results of a factorial experiment. *Tob. Control* 12, 178–183.
- Cruz, T.B., 2009. Monitoring the tobacco use epidemic IV. The vector: Tobacco industry data sources and recommendations for research and evaluation. *Prev. Med.* 48, S24–S34.
- Currivan, D.B., Nyman, A.L., Turner, C.F., Biener, L., 2004. Does telephone audio computer-assisted self-interviewing improve the accuracy of prevalence estimates of youth smoking? Evidence from the UMass Tobacco Study. *Public Opin. Q.* 68, 542–564.
- Delnevo, C.D., Gundersen, D.A., Hagman, B.T., 2008. Declining estimated prevalence of alcohol drinking and smoking among young adults nationally: artifacts of sample undercoverage? *Am. J. Epidemiol.* 167, 15–19.
- Delnevo, C.D., 2006. Smokers choice: what explains the steady rise of cigar use in the US? *Public Health Reports* 12 (2), 116–119.
- Dent, C.W., Galaif, J., Sussman, S., Stacy, A., Burton, D., Flay, B.R., 1993. Demographic, psychosocial and behavioral differences in samples of actively and passively consented adolescents. *Addict. Behav.* 18, 51–56.
- Eaton, D.K., Lowry, R., Brener, N.D., Grunbaum, J.A., Kann, L., 2004. Passive versus active parental permission in school-based survey research: does the type of permission affect prevalence estimates of risk behaviors? *Eval. Rev.* 28, 564–577.
- Farrelly, M.C., 2009. Monitoring the tobacco use epidemic V. The environment: Factors that influence tobacco use. *Prev. Med.* 48, S35–S43.
- Fong, G.T., Cummings, K.M., Borland, R., Hastings, G., Hyland, A., Giovino, G.A., Hammond, D., Thompson, M.E., 2006. The conceptual framework of the International Tobacco Control (ITC) Policy Evaluation Project. *Tob. Control* 15, iii3–iii11.
- Gfroerer J, Eyerma J, and Chromy J, Eds (2002) Redesigning an ongoing national household survey: Methodological issues. DHHS Publication No. SMA 03–3768. Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- Gfroerer, J., Wright, D., Kopstein, A., 1997. Prevalence of youth substance use: the impact of methodological differences between two national surveys. *Drug Alcohol Depend.* 47, 19–30.
- Giovino, G.A., Biener, L., Hartman, A.M., et al., 2009. Monitoring the tobacco use epidemic I. Overview: Optimizing measurement to facilitate change. *Prev. Med.* 48, S4–S10.
- Hox, J.J., de Leeuw, E., 1994. A comparison of nonresponse in mail, telephone, and face-to-face surveys. *Quality and Quantity: International Journal of Methodology* 28, 329–344.
- Jha, P., Chaloupka, F.J., 2000. Tobacco control in developing countries. Oxford University Press, Oxford, UK.

- Johnston, L.D., O'Malley, P.M., Bachman, J.G. and Schulenberg, J.E. (2005) Monitoring the Future national results on adolescent drug use: Overview of key findings, 2004. NIH pub no 05–5726. Bethesda, MD: National Institute on Drug Abuse.
- Kalsbeek, W., Agans, R., Pyles, A., Pevzner, E.S., Bowling, J., Ribisl, K., 2003. The influence of design on estimates of youth smoking behavior in school-based surveys. Presentation at 2003. Joint Statistical Meeting, San Francisco, CA.
- Kann, L., Brener, N.D., Warren, C.W., Collins, J.L., Giovino, G.A., 2002. An assessment of the effect of data collection setting on the prevalence of health risk behaviors among adolescents. *J. Adolesc. Health* 31, 327–335.
- Kennet, J., and Gfroerer, J. (2005) Evaluating and improving methods used in the National Survey on Drug use and Health (DHHS Publication No. SMA 05-4044, Methodology Series M-5). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.L., Walters, E.E., Zaslavsky, A.M., 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol. Med.* 32, 959–976.
- Klein, S. M., Giovino, G. A., Barker, D. C., Tworek, C., Cummings, K. M., O'Connor, R. J., and K. K. L., (in press) Use of Flavored Cigarettes among Older Adolescent and Adult Smokers: United States, 2004. *Nicotine and Tobacco Research*.
- National Cancer Institute, 1998. *Cigars: Health Effects and Trends*. Bethesda, Md: US Dept of Health and Human Services, Public Health Service. Smoking and Tobacco Control Monograph No. 9. NIH Publication 98-4302.
- National Cancer Institute (2008a) Tobacco Use Supplement- Current Population Survey. Available at: <http://riskfactor.cancer.gov/studies/tus-cps>. Accessed January 23, 2008.
- National Cancer Institute (2008b) Health Information National Trends Survey. Available at: <http://hints.cancer.gov>. Accessed January 23, 2008.
- O'Connor, R.J., Giovino, G.A., Kozlowski, L.T., Shiffman, S., Hyland, A., Bernert, J.T., Caraballo, R.S., Cummings, K.M., 2006. Changes in nicotine intake and cigarette use over time in two nationally representative cross-sectional samples of smokers. *Am. J. Epidemiol.* 164, 750–759.
- Pan, Y., and de la Puente, M. (2005) Census bureau guideline for the translation of data collection instruments and supporting materials: documentation on how the guideline was developed. U.S. Bureau of the Census, Statistical Research Division.
- Pirkle, J.L., Bernert, J.T., Caudill, S.P., Sosnoff, C.S., Pechacek, T.F., 2006. Trends in the exposure of nonsmokers in the U.S. population to secondhand smoke: 1988–2002. *Environ. Health Perspect.* 114, 853–858.
- Pokorny, S.B., Jason, L.A., Schoeny, M.E., Townsend, S.M., Curie, C.J., 2001. Do participation rates change when active consent procedures replace passive consent. *Eval. Rev.* 25, 567–580.
- Ramsey, L.T., Pelletier, A., Knight, S., 2004. Differences in smoking prevalence between the adult tobacco survey and the behavioral risk factor surveillance system. *Prev. Chronic. Dis.* 1, A22.
- RTI International (2006) Third Annual Independent Evaluation of New York's Tobacco Control Program Final Report, New York State Department of Health, September 2006, available at: [www.health.state.ny.us/prevention/tobacco\\_control/docs/independent\\_evaluation\\_report\\_2006.pdf](http://www.health.state.ny.us/prevention/tobacco_control/docs/independent_evaluation_report_2006.pdf). Accessed January 23, 2008.
- Siller, A. B., and Tompkin, L. (2005) The Big Four: Analyzing Complex Sample Survey Data Using SAS®, SPSS®, STATA®, and SUDAAN® 18th Annual Conference of the NorthEast SAS Users Group, September 11–14, 2005.
- Smith, E.A., Malone, R.E., 2003. The outing of Philip Morris: advertising tobacco to gay men. *Am. J. Public Health* 93, 988–993.
- Steeh, N., Kirgis, B., DeWitt, C.A.J., 2001. Are they really as bad as they seem? Nonresponse rates at the end of the twentieth century. *J. Official Stat.* 17, 227–247.
- Stellman, S.D., Djordjevic, M.V., 2009. Monitoring the tobacco use epidemic II. The agent: Current and emerging tobacco products. *Prev. Med.* 48, S11–S15.
- Substance Abuse and Mental Health Services Administration Results from the 2005. National Survey on Drug use and Health: National Findings (Office of Applied Studies, NSDUH Series H-30, DHHS Publication No. SMA 06-4194), 2006 Rockville, MD.
- U.S. Bureau of the Census (2003) Language Use and English-Speaking Ability: 2000 Census Brief Available at <http://www.census.gov/prod/2003pubs/c2kbr-29.pdf>.
- U.S. Department of Health and Human Services Tobacco Use Among U.S. Racial/Ethnic Minority Groups—African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: A Report of the Surgeon General. Atlanta, Georgia: 1998.
- University of Michigan Institute for Social Research. Panel Study on Income Dynamics (2008) Panel Study on Income Dynamics. Available at: <http://psidonline.isr.umich.edu/>. Accessed March 12, 2008.
- University of North Carolina—Carolina Population Center National Longitudinal Study of Adolescent Health (Add Health) (2008) Available at: <http://www.cpc.unc.edu/projects/addhealth>. Accessed March 12, 2008.
- US Department of Agriculture (2006) Tobacco outlook. Washington: U.S. Department of Agriculture, Economic Research Service. (TBS-260).
- Willis, G., Lawrence, D., Thompson, F., 2005. The use of cognitive interviewing to evaluate translated survey questions: lessons learned. Conference of the Federal Committee on Statistical Methodology.
- Yee, S.L., Schooley, M., 2001. Surveillance and Evaluation Data Resources for Comprehensive Tobacco Control Programs. Centers for Disease Control and Prevention, Atlanta(GA).