HHS Public Access

Author manuscript

Tob Control. Author manuscript; available in PMC 2024 October 20.

Published in final edited form as:

Tob Control. ; 33(6): 733–738. doi:10.1136/tc-2022-057851.

Update to the Design and Methods of the PATH Study, Wave 4 (2016–2017)

Jean D. Opsomer[1], **Sylvia Dohrmann**[1], **Ralph DiGaetano**[1], **Andrea Piesse**[1], **Ethel Taylor**[2], **MeLisa R. Creamer**[3], **Daifeng Han**[1], **Colm Everard**[3], **Nicolette Borek**[2], **Frost Hubbard**[1], **Izabella Zandberg**[2], **Moana Kanel**[3], **Kristie Taylor**[1], **Heather L. Kimmel**[3], **Antonio Paredes**[2], **Yumiko Siegfried**[1], **Yu-Ching Cheng**[2], **Jonathan Kwan**[2], **Charles Carusi**[1], **Andrew Hyland**[4]

[1]Westat, Rockville, Maryland, USA

[2]Center for Tobacco Products, US Food and Drug Administration, Department of Health and Human Services, Silver Spring, Maryland, USA

[3]National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA

[4]Roswell Park Cancer Institute, Buffalo, NY, USA

Abstract

The Population Assessment of Tobacco and Health (PATH) Study is a nationally representative study of the U.S. population on tobacco use and its effects on health, with four waves of data collection between 2013 and 2017. Prior work described the methods of the first three waves. In this paper, we describe the methods of Wave 4, during which a replenishment sample was added to the ongoing cohort. We describe the design and estimation methods of the continuing Wave 1 Cohort (with four waves of data) and the Wave 4 Cohort (the new cohort created at Wave 4). We provide survey quality metrics, including response rates for both cohorts and a panel conditioning analysis, and provide guidance on understanding the target populations for both cohorts.

INTRODUCTION

The Population Assessment of Tobacco and Health (PATH) Study is a nationally representative cohort study of the U.S. population on tobacco use and its effects on health. Data collected in the PATH Study are used for longitudinal and cross-sectional estimates that inform tobacco regulatory science, in particular regulatory actions of the Food and Drug Administration's Center for Tobacco Products. Researchers and policy makers outside

Correspondence to Dr. Opsomer, Westat, 1600 Research Boulevard, Rockville, MD 20850, USA jeanopsomer@westat.com.

Contributorship

JO: lead author, guarantor. SD, RD, AP, DH, FH, KT, YS, CC, AH: participated in planning, conducting and reporting the work. ET, MC, CE, NB, IZ, MK, HK, AP, YC, JK: participated in planning and reporting the work.

Conflicts of interest disclosures: The authors declare no conflicts of interest.

Financial disclosures. None of the authors have any financial disclosures to report.

Disclaimer:

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U.S. Department of Health and Human Services or any of its affiliated institutions or agencies.

the United States can learn from research conducted in the U.S. context (for example, association of flavored tobacco with later continued tobacco use[1], measures of dependence on different tobacco products[2]) and can build on the study's methodological experience. The range of health conditions and risk factors for physical and mental health covered in the PATH Study is diverse enough to be relevant to researchers both within and outside the field of tobacco regulatory science. For example, the PATH Study includes data on substance use and media use, as well as biomarker data relevant to certain health conditions. Various PATH Study survey measures have been included in the PhenX Toolkit[3], a web-based catalog of recommended measurement protocols selected by experts to include in studies with human participants.

Data collection for the PATH Study started in 2013. Interviews, conducted with tobacco users and nonusers ages 12 and older, provide data about a wide range of tobacco products. Biomarkers of exposure and potential harm are measured in urine and blood specimens from subsamples of respondents. Hyland et al.[4] describe the conceptual model of the study and the design and implementation of Wave 1, which involved recruitment of the initial cohort in 2013–2014. Additional waves of data were collected for this cohort, including Wave 2 in 2014–2015 and Wave 3 in 2015–2016, with the methods described in Piesse et al.[5]. This ongoing cohort is referred to as the Wave 1 Cohort. This article focuses on Wave 4 of the PATH Study, which continued data collection on the Wave 1 Cohort and added a replenishment sample of adults and youth in 2016–2017. The combined set of Wave 4 participants from the Wave 1 Cohort and the replenishment sample forms the Wave 4 Cohort.

Longitudinal surveys can be constructed using a variety of designs. One simple design option is to use a fixed panel representing its original population only. Such a panel can be supplemented with population "births," when new entrants into the population can be sampled separately. When a survey is intended to study an evolving population, it is more common to use series of rotating panels. These panels may overlap in time, and each panel is usually intended to represent the same population at different points in time. Cross-sectional and short-duration longitudinal estimates are then constructed by combining sampling units across panels. For a discussion of longitudinal sample designs, we refer to Buck et al.[6]. The PATH Study combines characteristics of several longitudinal designs, and could be referred to as a mixed panel design. It started as a fixed panel, the Wave 1 Cohort, but with the creation of the Wave 4 Cohort, it incorporates aspects of rotating panels because Wave 4 estimates are constructed by combining sampling units that were recruited in the sample at different times.

Longitudinal surveys provide data that enable estimating changes in the characteristics and behavior of individuals over time, which is typically not possible with a cross-sectional survey. However, maintaining the analytic usefulness of a longitudinal survey such as the PATH Study requires addressing three important challenges:

- Changes in sample size and composition over time: Due to sample attrition, the overall sample size decreases, resulting in less power to detect significant results. In particular, in addition to outright loss of sample members, the sample sizes

of population groups of interest defined by personal characteristics that change over time (e.g., narrowly defined age groups such as youth) can dwindle as participants naturally progress out of the group in which they were originally sampled.

- Panel conditioning[7]: Panel conditioning occurs when the reported behaviors of survey respondents who have been retained in a study for an extended period are influenced by their participation in the study itself.
- Cross-sectional representativeness: Because of gradual changes in the underlying population, the original sample becomes progressively less representative of the evolving population over time. This divergence in characteristics of the sample and the population can become sufficient to degrade the cross-sectional value of the survey estimates. This challenge only concerns longitudinal surveys that are also used for cross-sectional estimation, which is the case for the PATH Study.

As will be further described in this article, the addition of the replenishment sample to the PATH Study at Wave 4 is intended to address these challenges.

METHODS

PATH Study Sample Design Features

The Wave 1 target population consists of individuals ages 12 and older in the U.S. civilian noninstitutionalized population (CNP) at the time of Wave 1. A nationally representative household sample was selected at Wave 1 through a multi-stage design. At the first stage, a stratified sample of 156 geographical primary sampling units (PSUs) was selected. Each PSU consisted of one or several adjacent counties. Within the selected PSUs, smaller geographical areas, referred to as segments, were sampled at the second stage. Later stages sampled mailing addresses in the selected segments, with further sampling of individuals from households identified at these addresses through a brief in-person interview, known as the household screener. Also at Wave 1, a sample of individuals ages 9–11 was selected, with the goal of providing participants ages 12 and 13 in later waves. These individuals are referred to as “shadow youth.” More details on the Wave 1 PATH Study design are in Hyland et al.[4]. All individuals selected at Wave 1, including the shadow youth, make up the Wave 1 Cohort. After selection of the sample at Wave 1, the set of participants remained fixed at Wave 2 and Wave 3 except for individuals residing outside the United States or incarcerated who were not eligible at those waves. Because of the presence of shadow youth, the Wave 2 and Wave 3 target populations continued to consist of individuals ages 12 and older. Piesse et al.[5] provide details on Wave 2 and Wave 3 of the PATH Study.

At Wave 4, previous participants in the study were reinterviewed and an additional replenishment sample of new participants was selected. These were combined to form the Wave 4 Cohort of the study. In developing the sample design for the replenishment sample, multiple factors were considered: sample attrition, aging of the youth sample, observed tobacco use among study participants, time interval between the primary PATH Study waves, and time until the next planned replenishment wave. Specifically, there were to be 2 years between the primary waves after Wave 4, with the next replenishment sample

scheduled for Wave 6. Thus, only 10- and 11-year-olds were sampled at Wave 4 as shadow youth to cover the age range 12–13 at Wave 5. Since the replenishment sample was selected from members of the CNP at the time of Wave 4, only Wave 1 Cohort respondents who were in the CNP at Wave 4 were eligible to be in the Wave 4 Cohort. The replenishment sample provided an opportunity for new immigrants and others entering or rejoining the CNP after Wave 1 to enter the PATH Study sample, improving its cross-sectional representativeness.

As part of the Wave 4 replenishment effort, adults, youth, and shadow youth were sampled within the existing PATH Study sample segments from among the addresses not selected for Wave 1. A new sample of 174,273 mailing addresses was selected. To meet the needs for the Wave 4 Cohort shadow sample, a randomly selected subset of the sampled addresses (115,536) was screened solely to identify shadow youth. The remaining addresses (58,737) were screened for adults, youth, and shadow youth. Due to the large number of addresses selected solely for the shadow youth replenishment sample, the first step in data collection was to mail these addresses a brief paper screener designed to determine the presence of youth age 10 or 11 in the household. If a paper screener was returned reporting the absence of such youth, the household at that address was not contacted further. Households at sampled addresses that returned the paper screener and reported a youth age 10 or 11, as well as households that did not return the paper screener, were contacted for in-person household screening.

The Wave 4 within-household sampling procedures mirrored those used at Wave 1 (see Hyland et al.[4]), with sampling rates varying for adults by age, race, and tobacco-use status, and the definitions of tobacco use expanded to include users of all electronic nicotine products. The within-household sampling rates were designed to bring the Wave 4 Cohort adult and youth sample sizes up to a level commensurate with Wave 1 sample sizes by sampling domain, including those defined by tobacco use and race. As a result, 13,047 adults and 5,313 youth ages 12 to 17 were selected as part of the replenishment sample. Among those sampled, 9,804 completed an interview and were combined with Wave 1 Cohort interview respondents at Wave 4 in the CNP at that time. These interview respondents, together with the newly sampled shadow youth, form the Wave 4 Cohort. Table 1 shows the sample sizes of the Wave 1 Cohort and the Wave 4 Cohort. The responding sample sizes at Wave 4 were 38,816 for the Wave 1 Cohort and 48,437 for the Wave 4 Cohort, respectively. The overall cohort samples include Wave 1 Cohort participants who did not respond at Wave 4 but will continue to be contacted at future waves, as well as 4,294 shadow youth who will enter the Wave 4 Cohort at future waves, when they become 12 years old. As the table indicates, the addition of the replenishment sample resulted in a new cohort at Wave 4 that is similar in size to the original Wave 1 sample, while also renewing representation of the younger ages.

Weighting and Estimation

At Wave 4, two longitudinal weights are available for the Wave 1 Cohort: the all-waves weight and the single-wave weight. These weights are similar to those described in Piesse et al.[5] for Wave 2 and Wave 3 of the PATH Study. The Wave 4 all-waves weight was assigned to Wave 4 respondents in the Wave 1 Cohort who responded at all waves since recruitment,

including any participants selected as shadow youth at Wave 1 who completed an interview for waves in which they were old enough to do so and whose parents verified their information at each earlier wave in which they were not yet old enough to be interviewed. The Wave 4 single-wave weight was assigned to all Wave 4 respondents in the Wave 1 Cohort whether or not they responded at Wave 2 or Wave 3, and was designed for analyses using data from Wave 1 and Wave 4 only.

The two sets of longitudinal weights for the Wave 1 Cohort are intended for statistically representative estimation of longitudinal characteristics of the CNP at the time of Wave 1. In addition, there is a *cross-sectional* weight for all Wave 4 respondents in the Wave 4 Cohort, allowing for statistically representative estimation of characteristics of the CNP at the time of Wave 4. The Wave 4 cross-sectional weight was assigned to respondents in the Wave 4 replenishment sample and Wave 4 respondents from the Wave 1 Cohort who were in the CNP at Wave 4. The weighting process for the Wave 4 Cohort, further described below, differs from that used for previous waves, because it needs to account for the multiple ways a respondent can enter the study.

Weights for the replenishment sample respondents, which were selected from the Wave 4 CNP, were created in a manner similar to the original sample at Wave 1: the inverse of each sampled person's probability of selection multiplied by factors that adjust for nonresponse (see the Appendix of PATH Study[8] for a description of the weighting process). Most youth and adults selected for the replenishment sample at Wave 4 were eligible for selection at Wave 1 as they were in the CNP at that time, and most members of the Wave 1 Cohort were also members of the Wave 4 CNP. Questions were asked of the Wave 4 replenishment sample interview respondents to help determine retrospectively if they were members of the CNP at the time of Wave 1. If so, they had two chances of selection for the PATH Study. All other respondents had only a single chance of selection for the study. To account for this, the weights of the respondents with two chances of selection were adjusted and then combined, following an approach described by Watson[9]. This is an application of dual-frame estimation[10], a statistical method used to account for sampling partially overlapping populations referred to as *compositing*.

The Wave 4 single-wave weights for the Wave 1 Cohort members who were also members of the Wave 4 CNP were adjusted by a compositing factor $0 \leq \alpha \leq 1$, and those from the replenishment sample by a factor $(1 - \alpha)$ in compensation. Different values of α were applied for the respondents within subgroups corresponding to important sampling domains: four subgroups for adults, defined by the cross-classification of race (Black, non-Black) and age (ages 18–24, ages 25 and older); and six for youth, corresponding to the single years of age from 12 to 17. Within each subgroup, the effective sample sizes (i.e., the sample sizes divided by the design effects associated with the variation of the nonresponse-adjusted weights) were used to determine the value of α for the subgroup. Following compositing, the weights were calibrated via raking to external population totals from the 2016 American Community Survey (ACS) of the U.S. Census Bureau. The target population for the resulting weight is the CNP ages 12 and older at the time of Wave 4. For more information on the Wave 4 sample design and estimation methods, see PATH Study[10].

RESULTS

Response Rates

Table 2 presents weighted response rates for the Wave 1 Cohort and replenishment sample at Wave 4. In accordance with , weighted response rates for the Wave 1 Cohort and the replenishment sample were computed using inverse-of-probability-of-selection weights[11]. Further details on the Wave 4 response rate calculations appear in PATH Study[10].

The Wave 4 response rates condition on Wave 1 response for those sampled at Wave 1, whereas the person-level response rates for the replenishment sample at Wave 4 condition on completion of the Wave 4 household screener. Given these conceptual differences, no attempt was made to compute Wave 4 response rates for the Wave 4 Cohort as a whole.

PANEL CONDITIONING ANALYSIS

Cantor[12] describes three types of panel conditioning effects that can result from participation in a longitudinal survey, including (1) changes in behavior, (2) changes in reported behavior, and (3) changes in attitudes and opinions. Panel conditioning can appear in surveys on a wide range of topics, with the type and magnitude of the effects varying widely.

The Wave 4 replenishment sample, a new sample from the target population, provides an opportunity for a methodological assessment of potential panel conditioning in the PATH Study, and hence addresses the second challenge of longitudinal surveys mentioned in the Introduction. The assessment was based on the comparison of estimates from Wave 4 replenishment sample respondents with corresponding estimates developed for Wave 4 respondents who were sampled at Wave 1. For both samples, the estimates were computed using only members of the CNP at both Wave 1 and Wave 4 to maintain comparability. The weights used for the analysis were those established prior to the compositing of the two groups (these weights are not available to external analysts).

Table 3 shows the major subpopulations (youth, adults, young adults) and demographic subgroups within these subpopulations for which differences between the Wave 1 and replenishment sample estimates were investigated. A range of Wave 4 outcome measures was examined, including tobacco use, opinions about tobacco (youth only), and measures expected to be unaffected by panel conditioning to serve as “control” measures (social media use, health). Tables S1 and S2 in the Supplementary Materials show the complete descriptions of these outcome measures.

Table 4 shows the results for youth (ages 12–17). For the Wave 4 tobacco-use measures, there was no evidence of consistent differences in estimates from the Wave 1 sample and the replenishment sample. The only overall difference that was statistically significant was small, and the maximum number of statistically significant differences detected among the subgroup estimates for these outcomes was 2 out of 12. As expected, there were very few differences among the control measures. However, there was evidence of panel conditioning related to two aspects of susceptibility to specific tobacco products: curiosity about e-

products and hookah, and an interest in trying e-products or hookah soon. The replenishment sample estimates were higher for these measures, consistent with the hypothesis that a higher percentage of youth in the replenishment sample would be interested or curious, given that most Wave 1 youth would already have received exposure to new tobacco products over several waves of data collection by Wave 4. Analysts using curiosity variables in modeling should be aware of this effect and interpret the resulting estimates with some care.

Table 5 shows the results for adults and the young adult subset. For both groups, there was little evidence of panel conditioning for the current tobacco-use measures, with the statistical significance of otherwise small differences due to the very large sample sizes (see Table 1). However, for the two health-based control measures as well as the “ever e-product user” measure, the Wave 1 sample had consistently higher estimates. Using the Cantor categories, these differences could correspond to a change in actual behavior or to a change in reported behavior not corresponding to an actual change. The latter interpretation appears more likely given the lack of differences in the current use estimates and the fact that these differences were also observed for the control measures. PATH Study participants were asked at their first interview about “ever use” of tobacco products and whether they were “ever told” they had certain health conditions, and then asked only about these same items in reference to the past 12 months in follow-up interviews. There is reason to believe that the PATH Study estimates of “ever use” may be more accurate than similar estimates from cross-sectional studies, because combining successive recollections within past years is more accurate than a one-time recollection over a lifetime.

DISCUSSION

As noted in the Introduction, one of the challenges for longitudinal studies is an increasing discrepancy between the sample and the target population, which degrades the cross-sectional estimation capabilities of the panel. Here, we discuss this challenge in more detail. At Wave 1, estimates from the PATH Study allowed statistically representative estimation of characteristics of its target population, consisting of individuals ages 12 and older in the CNP. At Waves 2 and 3, the target population remains the Wave 1 CNP (further restricted to individuals who were in the United States and not incarcerated at the time of the later wave). Because the time between Wave 1 and Wave 3 is only 2 years, differences in their CNPs are expected to be small, and the shadow youth selected at Wave 1 ensure that 12- and 13-year-olds continued to appear in the sample. However, differences in the CNP increase with time, so the cross-sectional estimation capabilities of the PATH Study would degrade unless the sample is replenished.

By adding the replenishment sample at Wave 4, the overall PATH Study sample is “reset” in the sense that the Wave 4 Cohort (consisting of the ongoing cohort from Wave 1 plus the replenishment sample) is statistically representative for the CNP at Wave 4, and estimates for 12-and 13-year-olds can be produced at Wave 5. Combined with the first three waves of the PATH Study, the resulting data make it possible to obtain longitudinal estimates for any two waves between Wave 1 and Wave 4 for the Wave 1 CNP, and cross-sectional estimates for the CNP at Wave 1 and at Wave 4.

Until its next replenishment, the PATH Study will have two longitudinal cohorts: the Wave 1 Cohort, which is statistically representative for the CNP at Wave 1, and the Wave 4 Cohort, statistically representative for the CNP at Wave 4. The Wave 1 Cohort provides extensive data over a longer period of time for longitudinal analyses but because of aging up, has no participants younger than age 12 at Wave 4. The minimum age of participants will continue to increase and youth, an important subpopulation within the Wave 1 Cohort, will eventually transition to adults. The Wave 4 Cohort provides cross-sectional estimates reflective of the CNP at Wave 4, has its sample size restored to a similar level as Wave 1 originally, and has newly enrolled individuals as young as age 10 who will be interviewed in future waves. Additional replenishment waves are planned, to maintain the cross-sectional representation of the study and ensure that the younger ages continue to be represented. These replenishments will also be used to perform panel conditioning analyses at these waves, to continue to monitor the quality and representativeness of the sample data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This manuscript is supported with Federal funds from the National Institute on Drug Abuse (NIDA), National Institutes of Health (NIH), and the Center for Tobacco Products (CTP), Food and Drug Administration (FDA), Department of Health and Human Services, under contract to Westat (contract nos. HHSN271201100027C and HHSN271201600001C), and through an interagency agreement between NIH NIDA and FDA CTP. Dr. Kimmel and Dr. Creamer were substantially involved in the scientific management of and providing scientific expertise for contract nos. HHSN271201100027C and HHSN271201600001C. This article was prepared while Izabella Zandberg was employed at the Food and Drug Administration, Center for Tobacco Products.

REFERENCES

1. Villanti AC, Johnson AL, Glasser AM, et al. Association of flavored tobacco use with tobacco initiation and subsequent use Among US youth and adults, 2013–2015. JAMA Netw Open 2019;2(10):e1913804. doi:10.1001/jamanetworkopen.2019.13804. [PubMed: 31642927]
2. Strong DR, Leas E, Noble M, et al. Predictive validity of the adult tobacco dependence index: Findings from waves 1 and 2 of the Population Assessment of Tobacco and Health (PATH) study. Drug Alcohol Depend 2020;214:108134. doi:10.1016/j.drugalcdep.2020.108134. [PubMed: 32629146]
3. Hamilton CM, Strader LC, Pratt JG, et al. The PhenX Toolkit: Get the most from our measures. Am J Epidemiol 2011;174(3),253–60. doi: 10.1093/aje/kwr193. [PubMed: 21749974]
4. Hyland A, Ambrose BK, Conway KP, et al. Design and methods of the Population Assessment of Tobacco and Health (PATH) study. Tob Control 2017;26:371–378. doi:10.1136/tobaccocontrol-2016-052934. [PubMed: 27507901]
5. Piesse A, Opsomer J, Dohrmann S, et al. Longitudinal uses of the Population Assessment of Tobacco and Health study. Tobacco Regul Sci. 2021;7(1):3–16. doi: 10.18001/trs.7.1.1.
6. Buck N, Ermisch J, and Jenkins SP. Choosing a longitudinal survey design: the issues. Colchester, UK: ESRC Research Centre on Micro-Social Change, University of Essex; Sept. 1995. 60 p.
7. Cantwell PJ. Panel conditioning. In Lavrakas P, ed. Encyclopedia of Survey Research Methods. Sage Publications, Thousand Oaks, CA 2008: 566–567.
8. United States Department of Health and Human Services. National Institutes of Health. National Institute on Drug Abuse, and United States Department of Health and Human Services. Food and Drug Administration. Center for Tobacco Products. Population Assessment of Tobacco and Health (PATH) Study [United States] Restricted-Use Files, User Guide. ICPSR36231-v30 Ann Arbor,

MI: Inter-university Consortium for Political and Social Research [distributor], 2022-4-21. 10.3886/ICPSR36231.userguide

9. Watson N Evaluation of weighting methods to integrate a top-up sample with an ongoing longitudinal sample. Surv Res Methods 2014; [S.l.], 8 (3), 195–208, ISSN 1864–3361. 10.18148/srm/2014.v8i3.5818
10. Lohr SL and Rao JNK. Inference for dual frame surveys. J Am Stat Assoc 2000: 95:271–280. doi: 10.1080/01621459.2000.10473920.
11. American Association of Public Opinion Research (AAPOR). Standard definitions: Final dispositions of case codes and outcome rates for surveys. 9th ed. Oakbridge Terrace, IL: AAPOR 2016. https://www.aapor.org/aapor_main/media/publications/standard-definitions20169theditionfinal.pdf. Date accessed 04 April 2022.
12. Cantor D A review and summary of studies on panel conditioning. In Menard S, ed. Handbook of Longitudinal Research: Design, Measurement, and Analysis Across the Social Sciences. Elsevier, Burlington, VT: 2008: 123–138.

What is already known on this topic –

The PATH Study is an ongoing representative study of tobacco use and health in the United States.

What this study adds –

Wave 4 of the PATH Study adds a replenishment sample to the existing cohort, increasing its analytic capabilities but making its structure more complex. This study explains the new structure and provides associated quality measures.

How this study might affect research, practice, or policy –

This study will assist researchers and policy makers in the interpretation of results based on PATH Study data.

Table 1.

Wave 4 sample sizes of the Wave 1 Cohort and Wave 4 Cohorts of the Population Assessment of Tobacco and Health (PATH) Study.

	Sample	Wave 1 Cohort	Wave 4 Cohort
Sampled at Wave 1	Wave 1 Cohort members who did not complete a Wave 4 interview	14,362	-
	Wave 4 interview respondents in the Wave 1 Cohort and in the Wave 4 CNP [a]	*38,633*	*38,633*
	Wave 4 interview respondents in the Wave 1 Cohort, but not in the Wave 4 CNP [a]	*183*	-
Sampled at Wave 4	*Wave 4 replenishment sample interview respondents*	-	*9,804*
	Wave 4 replenishment sample shadow youth respondents (not interviewed at Wave 4, will provide participants at later waves)	-	4,294
Wave 4 respondent sample size		*38,816*	*48,437*
Total sample size (shadow youth, youth and adults)		53,178	52,731

[a]Civilian noninstitutionalized population

Table 2.

Population Assessment of Tobacco and Health (PATH) Study response rates at Wave 4 (2016–2017).

		Weighted response rate (%)
Replenishment sample	Household screener	51.2
	Adult interview[a]	68.0
	Youth interview[a]	70.6
	Shadow youth parental consent[a]	78.7
Wave 1 Cohort	Adult interview[b]	73.5
	Youth interview[b]	79.5

[a]The replenishment sample interview and parental consent response rates condition on household screener response.

[b]The Wave 1 sample response rates condition on Wave 1 response.

Table 3.

Wave 4 subgroups used for panel conditioning assessment, Population Assessment of Tobacco and Health (PATH) Study.

Demographic	Youth (12–17)	Adult (18+)	Young adult (18–24)
Age	Ages 12–15	Ages 18–24	Age 18
	Ages 16–17	Ages 25+	Ages 19–20
			Ages 21–22
			Ages 23–24
Sex	Male	Male	Male
	Female	Female	Female
Race/Ethnicity	Hispanic	Hispanic	Hispanic
	White alone, non-Hispanic	White alone, non-Hispanic	White alone, non-Hispanic
	Black alone, non-Hispanic	Black alone, non-Hispanic	Black alone, non-Hispanic
	Other race or multiple races, non-Hispanic	Other race or multiple races, non-Hispanic	Other race or multiple races, non-Hispanic
Education		Less than high school	
		GED	
		High school	
		Some college but no degree, or associate’s degree	
		Bachelor’s degree	
		Advanced degree	
Work Status		Full-time work (35+ hours per week)	Full-time work (35+ hours per week)
		Part-time work (< 35 hours per week)	Part-time work (< 35 hours per week)
		Not working for pay	Not working for pay
Region	Northeast	Northeast	Northeast
	Midwest	Midwest	Midwest
	South	South	South
	West	West	West

Table 4.

Comparison of weighted estimates for Wave 4 Cohort youth, by recruitment wave[a], Population Assessment of Tobacco and Health (PATH) Study.

Survey question	Estimate for those sampled at Wave 1 (%)	Estimate for those sampled at Wave 4 (%)	Difference[b]	Number of significant differences
Current cigarette user	3.2	3.0	−0.1	0
Current E-product user	4.3	4.3	0.0	2
Current any tobacco user	7.6	7.0	−0.6	3
Ever cigarette user	9.9	11.0	1.1	1
Ever E-product user	15.8	14.1	−1.7*	1
Thought of chemicals often	8.5	9.2	0.7	2
Some tobacco safer	25.5	25.8	0.3	0
Curious about E-products	19.4	25.6	6.2*	11
Curious about hookah	14.6	22.5	7.9*	12
Try E-product soon	11.6	14.1	2.5*	5
Try hookah soon	9.3	12.9	3.7*	9
Social Media	86.7	87.9	1.2	3
Health good/excellent	86.7	86.9	0.2	0

[a]Estimates are shown for all youth, as well as the counts of the statistically significant differences for the 12 subgroups in Table 3.

[b]Significant difference at the .05 level denoted by *.

Table 5.

Comparison of weighted estimates for Wave 4 Cohort adults (18+) and young adults (18–24) by recruitment wave [a], Population Assessment of Tobacco and Health (PATH) Study.

Survey question	Estimate for those sampled at Wave 1 (%)	Estimate for those sampled at Wave 4 (%)	Difference [b]	Number of significant differences
All adults				
Current cigarette user	18.0	16.1	−1.9*	9
Current E-product user	3.3	3.0	−0.3	2
Current any tobacco user	23.0	21.0	−1.9*	4
Ever cigarette user	68.1	65.8	−2.3	0
Ever E-product user	28.3	23.1	−5.2*	18
Ever high blood pressure	35.6	29.0	−6.6*	14
Ever lung problem	25.2	18.3	−6.9*	17
Young adult subset				
Current cigarette user	14.7	13.9	−0.9	2
Current E-product user	6.7	6.7	0.0	1
Current any tobacco user	22.8	22.0	−0.7	1
Ever cigarette user	49.2	49.0	−0.2	1
Ever E-product user	53.1	45.9	−7.2*	12
Ever high blood pressure	7.5	4.5	−3.0*	14
Ever lung problem	25.8	19.7	−6.1*	13

[a]Estimates are shown for all adults/young adults, as well as the counts of the significant differences for the subgroups in Table 3. The adult analysis included 21 subgroups and the young adult analysis included 17 subgroups.

[b]Significant difference at the .05 level denoted by *.