

Attachment 20

**PATH Study Interim
Report**

Population Assessment of Tobacco and Health Study

Interim Report to the Office of
Management and Budget on Data
and Biospecimen Collection from
Waves 1, 2, and 3

June 24, 2016

Submitted by:

Kevin P. Conway, Ph.D.

Deputy Director

Division of Epidemiology, Services, and Prevention Research

National Institute on Drug Abuse

6001 Executive Blvd., Room 5185

Rockville, MD 20852

Phone: 301-443-8755

Table of Contents

<u>Chapter</u>		<u>Page</u>
1	Introduction	1
1.1	Purpose of Interim Report (Terms of Clearance).....	1
1.2	Sample Design.....	2
1.2.1	Overview of Sample Design for Wave 1 (Baseline Wave).....	2
1.2.2	Overview of Sample Design for Waves 2 and 3	3
1.3	Summary of Wave 1 Findings	5
1.3.1	Wave 1 Response Rates	5
1.3.2	Wave 1 Nonresponse Bias Analysis.....	6
1.3.3	Wave 1 Statistical Approach for Addressing Nonresponse	7
2	Wave 2.....	9
2.1	Response Rates.....	9
2.1.1	Retention Rates for Continuing Adults and Continuing Youth.....	11
2.1.2	Recruitment Rates for Aged-up Adults and Aged-up Youth	15
2.1.3	Biospecimen Collections	19
2.2	Nonresponse Bias Analysis	22
2.2.1	Method	23
2.2.2	Results	25
2.3	Statistical Approach for Addressing Nonresponse.....	40
2.3.1	Interviews.....	40
2.3.2	Biospecimens.....	56
2.4	Summary of Findings	56

<u>Chapter</u>		<u>Page</u>
3	Wave 3.....	59
	3.1 Predicted Response Rates.....	60
	3.1.1 Predicted Retention Rates among Wave 2 Respondents for Continuing Adults and Continuing Youth.....	61
	3.1.2 Predicted Recruitment Rates among Wave 2 Respondents for Aged-up Adults and Aged-up Youth.....	65
	3.1.3 Predicted Retention and Recruitment Rates among Wave 1 Respondents.....	69
	3.1.4 Biospecimen Collections	70
	3.2 Nonresponse Bias Analysis	73
	3.3 Statistical Approach for Addressing Nonresponse.....	82
	3.4 Summary of Findings	83
4	Conclusions and Implications for Study Going Forward.....	85
	References	89
 <u>Appendix</u>		
A	Cigarette Smoking Questions in the PATH Study and Other Surveys	A-1
 <u>Table</u>		
1-1	Summary of PATH Study Wave 1 response rates.....	6
2-1	Case counts by Wave 1 adult/youth/shadow youth classification and Wave 2 adult/youth/shadow youth classification.....	10
2-2	PATH Study Wave 2 retention rates by Wave 1 characteristics: Adult Interview (continuing adults).....	13

<u>Table</u>		<u>Page</u>
2-3	PATH Study Wave 2 retention rates by Wave 1 characteristics: Youth Interview (continuing youth).....	14
2-4	PATH Study Wave 2 recruitment rates by Wave 1 characteristics: Adult Interview (aged-up adults).....	18
2-5	PATH Study Wave 2 recruitment rates by Wave 1 characteristics: Youth Interview (aged-up youth).....	18
2-6	PATH Study Wave 2 response rates by respondent characteristics: Urine collection (continuing adults).....	20
2-7	PATH Study Wave 2 response rates by respondent characteristics: Biospecimen collections (aged-up adults)	21
2-8	Comparison of Wave 2 continuing Adult Interview respondents with nonrespondents.....	26
2-9	Comparison of Wave 1 tobacco use rates for Wave 2 continuing Adult Interview respondents with nonrespondents	28
2-10	Comparison of Wave 2 continuing Youth Interview respondents with nonrespondents.....	29
2-11	Comparison of Wave 1 tobacco use rates for Wave 2 continuing Youth Interview respondents with nonrespondents.....	30
2-12	Comparison of Wave 2 aged-up Adult Interview respondents with nonrespondents.....	31
2-13	Comparison of Wave 1 tobacco use rates for Wave 2 aged-up Adult Interview respondents with nonrespondents	32
2-14	Comparison of Wave 2 aged-up Youth Interview respondents with nonrespondents.....	33
2-15	Comparison of Wave 2 aged-up adult urine specimen providers with non-providers of urine	35
2-16	Comparison of Wave 2 aged-up adult blood specimen providers with non-providers of blood.....	36

<u>Table</u>	<u>Page</u>
2-17 Current cigarette smoking based on Wave 2 Adult Interview respondents: IPS weights.....	37
2-18 Cigarette smoking based on Wave 2 Youth Interview respondents: IPS weights.....	38
2-19 Comparison of Wave 2 continuing Adult Interview respondents with Wave 1 Adult Interview respondents	45
2-20 Comparison of Wave 1 tobacco use rates for Wave 2 continuing Adult Interview respondents with Wave 1 Adult Interview respondents	47
2-21 Comparison of Wave 2 continuing Youth Interview respondents with Wave 1 youth respondents who were continuing youth at Wave 2.....	48
2-22 Comparison of Wave 1 tobacco use rates for Wave 2 continuing Youth Interview respondents with Wave 1 youth respondents who were continuing youth at Wave 2.....	49
2-23 Comparison of Wave 2 aged-up Adult Interview respondents with Wave 1 youth respondents who were aged-up adults at Wave 2.....	50
2-24 Comparison of Wave 1 tobacco use rates for Wave 2 aged-up Adult Interview respondents with Wave 1 youth respondents who were aged-up adults at Wave 2.....	51
2-25 Comparison of Wave 2 aged-up Youth Interview respondents with Wave 1 shadow youth who were aged-up youth at Wave 2.....	52
2-26 Current cigarette smoking based on Wave 2 Adult Interview respondents	54
2-27 Cigarette smoking based on Wave 2 Youth Interview respondents	55
2-28 Summary of PATH Study Wave 2 response rates.....	56

<u>Table</u>	<u>Page</u>	
3-1	Status of Wave 3 released cases from replicate group 1, as of April 29, 2016.....	60
3-2	PATH Study Wave 3 predicted retention rates by respondent characteristics: Adult Interview (continuing adults)	64
3-3	PATH Study Wave 3 predicted retention rates by respondent characteristics: Youth Interview (continuing youth)	65
3-4	PATH Study Wave 3 predicted recruitment rates by respondent characteristics: Adult Interview (aged-up adults)	68
3-5	PATH Study Wave 3 predicted recruitment rates by respondent characteristics: Youth Interview (aged-up youth)	68
3-6	PATH Study Wave 3 predicted retention and recruitment rates among Wave 1 respondents	69
3-7	PATH Study Wave 3 response rates by respondent characteristics: Urine collection (continuing adults)	71
3-8	PATH Study Wave 3 response rates by respondent characteristics: Biospecimen collections (aged-up adults)	72
3-9	Comparison of Wave 3 Adult Interview respondents with finalized and provisional nonrespondents (Wave 1 adults).....	75
3-10	Comparison of Wave 1 tobacco use rates for Wave 3 Adult Interview respondents with finalized and provisional nonrespondents (Wave 1 adults)	77
3-11	Comparison of Wave 3 interview respondents with finalized and provisional nonrespondents (Wave 1 youth).....	78
3-12	Comparison of Wave 1 tobacco use rates for Wave 3 respondents with finalized and provisional nonrespondents (Wave 1 youth)	79
3-13	Comparison of Wave 3 Youth Interview respondents with provisional nonrespondents (Wave 1 shadow youth)	81

<u>Table</u>		<u>Page</u>
3-14	Summary of PATH Study predicted response rates for Wave 3.....	83
A-1	Questions used to define adult current cigarette smoking in the PATH Study, TUS-CPS, NHIS, NHANES, and NSDUH	A-3
A-2	Questions used to define youth cigarette smoking in the PATH Study, NHANES, NSDUH, and NYTS	A-5

Introduction

1

The National Institute on Drug Abuse (NIDA), in partnership with the Food and Drug Administration (FDA), has prepared this report in response to the terms of clearance by the Office of Management and Budget (OMB) in its approval of Wave 3 of the Population Assessment of Tobacco and Health (PATH) Study.

1.1 Purpose of Interim Report (Terms of Clearance)

OMB approved the PATH Study's Revision Request for Wave 3 on August 17, 2015 (0925-0664). The terms of clearance of OMB's approval state: "Before submitting the information collection request for Wave 4 to OMB, NIDA/FDA should report to OMB regarding: a) the response rates associated with the full baseline wave [and full Wave 2], including screening, interview completion, and bio-specimen response; b) Wave 3 retention, recruitment rates for the "age in to adult" and "age in of shadow" subsamples; c) the results of nonresponse analysis and statistical approach for addressing non-response, as well as implications for the study going forward; and d) the statistical approach to be applied to the bio-specimen data to address potential non-response bias from lower consent and cooperation rates with this aspect of the study."

This report is organized in sections that correspond to OMB's terms of clearance: Section 1 includes a summary of findings from the 2015 Interim Report to OMB on the full Wave 1 response rates, nonresponse analysis, and statistical approach for addressing nonresponse. Section 2 presents the full Wave 2 response rates (retention and recruitment rates), results of a nonresponse analysis, and statistical approach for addressing nonresponse. Section 3 presents the predicted Wave 3 response rates (retention and recruitment rates), results of an interim nonresponse analysis, and statistical approach for addressing nonresponse. Section 4 summarizes the findings and considers their implications.

The report covers the full Waves 1 and 2 of the PATH Study, and data collected from a probability subsample of Wave 3 between October 19, 2015 and April 29, 2016. Response rates for Wave 2 and Wave 3 are compared throughout this report with corresponding rates projected in the PATH Study's Revision Requests to OMB for Wave 2 and Wave 3, respectively.

1.2 Sample Design

This section provides an overview of the sample design for the PATH Study and a description of replicate group 1; results in this report for Wave 3 are based on replicate group 1. Information on the study background and overall design is provided in Supporting Statement A of the PATH Study’s Revision Request to OMB for Wave 3.

1.2.1 Overview of Sample Design for Wave 1 (Baseline Wave)

The target population of the PATH Study at Wave 1 was the civilian, non-institutionalized U.S. population (i.e., including the 50 states and the District of Columbia) 12 years of age and older at that point in time. Thus, active duty military personnel and those residing in an institutional setting were excluded. College students living away from home during the school year were identified as members of their permanent residence (e.g., parents’ home). For Wave 1, a four-stage stratified area probability sample design was used with a two-phase design for sampling the adult cohort at the final stage. The sampling rates for adults varied by age, race, and tobacco use status. At the first stage, a stratified sample of geographical primary sampling units (PSUs) was selected, in which a PSU was a county or group of counties. For the second stage, within each selected PSU, smaller geographical segments (consisting of one or more census blocks) were formed and then a sample of these segments was drawn. At the third stage, a sample of addresses within sampled segments was drawn from listings of addresses; the main source of these addresses was obtained from the Postal Service (USPS) Computerized Delivery Sequence Files (CDSFs). The CDSFs provide very high coverage of the residential addresses in the U.S.

The fourth stage was the random selection of persons within the sampled households. A roster of all the members in the sampled household was constructed using the Household Screener. An adult household member, the household screener respondent, was asked to list members of the household and provide demographic as well as, for each adult, tobacco use information. This information was used in sampling three main groups of interest:

- Adults (up to two adults per household);
- Children ages 12 to 17 (referred to as “youth,” generally up to two per household); and

- Children ages 9 to 11 (referred to as “shadow youth,” generally up to two per household) to be enrolled in the youth cohort in later waves of the study on reaching 12 years of age.

Two-phase sampling was used for adult selection due to potential misreporting by the household screener respondent of the tobacco use status of other adult household members. The Phase 1 sampling depended on the age, race, and tobacco use information obtained from the Household Screener. The Phase 2 sampling was based on self-reported age, race, and tobacco use status, obtained by interviewing the individuals sampled at Phase 1. The sampling rates for the two phases were designed to achieve large enough sample sizes for young adults (ages 18 to 24) and adult tobacco users of all ages.

Because the full sample was selected using probability sampling methods, it is representative of the U.S. civilian non-institutionalized population 12 years of age and older. The PATH Study Wave 1 sample was divided into four replicate groups, consisting of probability samples of approximately 20 percent, 30 percent, 30 percent, and 20 percent of the sampled segments, respectively, within each sampled PSU. Each separate replicate group was a probability sample from the set of segments in the frame and, therefore, also representative of the civilian non-institutionalized U.S. population. The replicate groups were released to the field in a sequential manner (replicate group 1 in September 2013, replicate group 2 in November 2013, replicate group 3 in February and March 2014, and replicate group 4 in May 2014).

The PATH Study completed 32,320 Adult Interviews and 13,651 Youth Interviews in Wave 1. All adult interview respondents were asked to provide urine and blood specimens; 21,801 provided a urine specimen and 14,520 provided a blood specimen.

1.2.2 Overview of Sample Design for Waves 2 and 3

Wave 2 of the PATH Study was the first follow-up wave for participants in Wave 1. The target population for Wave 2 was the Wave 1 target population residing in the U.S. at Wave 2 with the exception of those who were incarcerated (with a corresponding definition for the target population at Wave 3). Thus, Wave 1 respondents who later joined the military or entered a health care institution (e.g., nursing home) were members of the target population and eligible for data collection for the PATH Study. At Wave 2 only Wave 1 respondents who died, resided outside the U.S., or were in a correctional facility were ineligible for a Wave 2 interview (and similarly for Wave 3).

Attempts were made to contact the Wave 1 youth and adult respondents as well as members of the shadow youth sample established at Wave 1. Youth from the shadow youth sample who turned age 12 by Wave 2 (aged-up youth) and were permitted by a parent or guardian to participate in the study were asked for assent to be interviewed for the first time at Wave 2. Similarly, persons in the youth sample at Wave 1 who reached age 18 by Wave 2 (aged-up adults) were asked to complete the adult instrument as well as to provide urine and blood specimens.

The PATH Study completed 28,375 Adult Interviews and 12,172 Youth Interviews in Wave 2. The study subsampled 14,465 adults for urine collection from adults who provided urine at Wave 1; among these subsampled adults, 12,569 completed the Wave 2 interview and 12,113 provided a urine specimen again at Wave 2. The study also collected urine and blood specimens from consenting aged-up adults (1,587 urine specimens and 908 blood specimens).

The PATH Study is currently conducting Wave 3, the second follow-up wave for participants in Wave 1. In addition, youth from the shadow youth sample established at Wave 1 who turn age 12 by Wave 3 and who are permitted by a parent or guardian to participate in the study are asked for assent to be interviewed for the first time at Wave 3. Similarly, persons in the youth sample at Wave 1 who reach age 18 by Wave 3 are asked to complete the adult instrument as well as to provide urine and blood specimens for the first time. The same follow-up rules apply for Wave 3 as for Wave 2:

- Wave 1 respondents who reside in the U.S. and are not incarcerated at Wave 3 are eligible for Wave 3 data collection, including those in the military or living in an institution at Wave 3; and
- In addition to Wave 2 respondents, Wave 2 nonrespondents are fielded for Wave 3 data collection, unless the nonresponse at Wave 2 was due to a firm or hostile refusal, inability to complete the Wave 2 interview in English or Spanish, death, or a physical or mental disability that prevents participation in the study.

For Wave 3, the PATH Study is subsampling 16,138 adults for urine collection from adults who provided urine at a previous wave. Among these adults, an estimated 12,000 are expected to provide a urine specimen again at Wave 3. The study is also collecting urine and blood specimens from consenting aged-up adults.

This Interim Report reviews the results for Wave 3 based on data collected from replicate group 1 through April 29, 2016. Wave 3 cases are released for follow up in monthly groups with the goal of

completing the Wave 3 interview as close as possible to the one-year anniversary date of the Wave 2 interview (or to the two-year anniversary date of the Wave 1 interview if no interview was obtained in Wave 2). Cases are fielded at the beginning of the calendar month prior to the month in which the anniversary date falls.¹ If there is more than one sampled person in the same household, the cases from that household are typically clustered for simultaneous release based on the earliest date of a completed interview for any sampled adult or youth in the household in Wave 2 (or in Wave 1, if nobody in the household completed a Wave 2 interview).² Thus, while the Wave 3 sample release is not explicitly tied to the release of the Wave 1 replicate groups, it does correlate with the distribution of completion dates for Wave 1 interviews. That is, a high percentage of cases in the first Wave 3 release groups came from replicate group 1, and replicate group 1 has higher percentages of released and finalized cases than the other replicate groups.

1.3 Summary of Wave 1 Findings

The 2015 Interim Report to OMB provided detailed findings on the full Wave 1 response rates, nonresponse analysis, and statistical approach for addressing nonresponse. This section summarizes those findings.

1.3.1 Wave 1 Response Rates

Sections 2.2 and 2.3 of the 2015 Interim Report describe the weight construction for Wave 1 of the PATH Study, and a brief summary also appears in Section 2.3.1 of this report. Inverse probability of selection (IPS) weights were used to compute the Wave 1 weighted response rates and to examine nonresponse bias. The final raked weights from Wave 1 were then used to demonstrate the effectiveness of the nonresponse weighting adjustments in reducing potential nonresponse bias.

¹ However, cases with anniversary dates in September through December 2016 (as determined by Wave 2 or Wave 1 interview dates) will all be released in the beginning of August, because Wave 3 data collection will end on October 31, 2016. Similarly, cases with anniversary dates in September through November 2015 were all released for data collection in October 2015 and were available for data collection beginning on the first day of Wave 3 data collection, October 19, 2015.

² If the gap between interviews within the household is greater than one month, interviewers are instructed to hold data collection for a sampled person with a later interview to a date closer to the individual's anniversary date whenever feasible and the interview would not be lost to nonresponse.

The weighted response rates for the PATH Study Household Screener, Adult and Youth Interviews, and biospecimen collections in Wave 1 are provided in Table 1-1.

Table 1-1. Summary of PATH Study Wave 1 response rates

Collection	Unweighted response rate, based on full Wave 1 sample	Weighted response rate, based on full Wave 1 sample
Household Screener	54.1%	54.0%
Adult Interview	74.8%	74.0%
Youth Interview	78.2%	78.4%
Urine	67.5%	63.6%
Blood	44.9%	43.0%

Differences in weighted response rates were modest for tobacco use status and demographic subgroups. The largest differential weighted response rate, 11.5 percentage points, was for the age of adults who provided urine specimens, which suggests a heightened potential for nonresponse bias. Notably, the differential weighted response rates for blood collection, ranging from 3.2 percentage points for ethnicity to 5.8 percentage points for race, were more consistent with those of other PATH Study collections at Wave 1.

1.3.2 Wave 1 Nonresponse Bias Analysis

A nonresponse bias analysis indicated that estimates of key demographic and tobacco use variables calculated from the PATH Study Wave 1 sample with the inverse probability of selection weights were comparable to those produced by national cross-sectional surveys. However, the completed household screening interviews from the Wave 1 sample appeared to underrepresent single- and two-person households relative to the 1-year 2013 American Community Survey (ACS) counts. The estimated percentage of persons who were non-Black and 25 years of age or older, from the household rosters, was also smaller than the corresponding estimate from the ACS.

Estimated distributions of demographic characteristics for adults completing the Adult Interview were similar to those from the 1-year 2013 ACS for race (except for persons in the “other race” category). Persons in the “other race” category were also underrepresented among the persons providing blood or urine specimens. The percentage of adults who were Hispanic estimated from those who provided blood specimens was similar to ACS population estimates, but, compared to ACS, Hispanics were overrepresented among adults who responded to the Adult Interview and those who provided urine specimens. In addition, the percentages of adults who were between 18

and 24 years old or between 25 and 44 years old as estimated from PATH Study interview respondents and from those who provided urine specimens were higher than the ACS estimates for these age groups. Males were underrepresented among respondents to the Adult Interview, and also among the persons who provided blood or urine specimens.

When compared to national cross-sectional surveys that measure adult tobacco use [the Tobacco Use Supplement to the Current Population Survey (TUS-CPS), the National Health Interview Survey (NHIS), the National Health and Nutrition Examination Survey (NHANES), and the National Survey on Drug Use and Health (NSDUH)], estimates of adult cigarette smoking from the PATH Study Wave 1 sample were roughly in the middle of the range of estimates for cigarette smoking. There is no indication of nonresponse bias with respect to this measure.

Estimates of demographic characteristics of youth in Wave 1 aligned with the 1-year 2013 ACS for most demographic characteristics. However, the estimated percentage of youth who were Hispanic from the PATH Study was significantly higher than the corresponding percentage estimated from the ACS.

PATH Study estimates of the selected youth cigarette smoking measure from the full Wave 1 sample were at the low end of estimates in comparison with national cross-sectional surveys that measure youth tobacco use (NHANES, NSDUH, and the National Youth Tobacco Survey (NYTS)). However, estimates based on these comparison surveys were from 2011 through 2013 while those based on the PATH Study were from September 2013 through December 2014, and evidence suggests that the use of traditional cigarettes is declining among youth. Although the difference in the time periods of the surveys is not by itself large enough to account for the different estimates, it is one of a number of factors that may explain the different estimates.

1.3.3 Wave 1 Statistical Approach for Addressing Nonresponse

The approach used to reduce potential nonresponse bias in the PATH Study is to adjust the weights of respondents at the household, adult, and youth levels to account for nonrespondents. Results of applying this approach to the full Wave 1 sample indicated the nonresponse adjustments were successful for reducing the discrepancy between the PATH Study estimates and 1-year estimates from the 2013 ACS with respect to demographic characteristics. Raked weights used for adults responding to the Adult Interview reduced differences between the PATH Study and ACS for adults providing biospecimens as well, for sex and ethnicity. The raking did not reduce differences in the

age distributions for the persons providing blood specimens, however. Additional sets of nonresponse-adjusted weights were therefore created for the persons who provided biospecimens, separately for urine and blood. Further biospecimen weights were created for the sets of person for whom urine and/or blood specimens were sent for laboratory analysis, using the procedure described in the memo on Wave 1 biospecimen weighting procedures submitted to OMB and approved on October 9, 2015.

Estimates of adult cigarette smoking using the IPS weights (before nonresponse adjustment) were in line with estimates from other surveys; agreement in these estimates was preserved using the nonresponse-adjusted weights. Weighting adjustments for youth corrected for the slight overestimate of the percentage of Hispanics among youth in Wave 1 but had little effect on the other demographic characteristics (i.e., IPS-weighted estimates already agreed with the ACS values) and estimates of youth cigarette smoking.

Wave 2

The PATH Study completed the Wave 2 data and biospecimen collections in October 2015.³ This section presents findings on the response rates for Wave 2, on the nonresponse analysis, and on the study's statistical approach for addressing nonresponse.

2.1 Response Rates

This section summarizes the retention and recruitment rates of the PATH Study for Wave 2, as well as the response rates for the biospecimen collections. The PATH Study Wave 2 Adult Interview and Youth Interview collected extensive self-report information through in-person data collection using audio computer-assisted self-interviewing (ACASI). The Adult Interview gathered information from adults (18 years old and older) about tobacco use behaviors, attitudes, knowledge, and health conditions, as well as information on demographics, environmental factors, family and peer influences, substance use, and general physical and mental health status. The Youth Interview gathered information from youth (12 to 17 years old) on topics similar to those in the Adult Interview. Youth were asked about their tobacco use and attitudes about tobacco, and for information on demographics, environmental factors, family and peer influences, substance use, and mental health. The PATH Study completed 28,375 Adult Interviews and 12,172 Youth Interviews in Wave 2.

Retention rates for Wave 2 apply to persons who completed the Adult Extended Interview in Wave 1 and persons who completed the Youth Extended Interview in Wave 1 and who were age 17 or younger at Wave 2. Recruitment rates for Wave 2 apply to those who had aged up, either as shadow youth who had turned age 12 and were eligible for the Wave 2 Youth Interview, or as youth in Wave 1 who had turned age 18, thus becoming eligible to participate as adults in Wave 2. Specifically,

- All the adults from Wave 1 are continuing adults and, if eligible, were asked to complete an Adult Interview in Wave 2.

³ A small number of blood collections for Wave 2 were completed in November 2015.

- The youth from Wave 1 who completed a Youth Interview at Wave 2 are continuing youth, and those who completed an Adult Interview at Wave 2 are aged-up adults.
- The Wave 1 shadow youth who completed the Wave 2 Youth Interview are aged-up youth.

A Wave 1 participant who was a nonrespondent at Wave 2, however, does not have a Wave 2 interview date, so his/her Wave 2 age was determined using the date of birth or age information collected in Wave 1 as well as any updated information obtained between Waves 1 and 2. Each Wave 1 respondent had an “anniversary month” for Wave 2, which was approximately one year after the Wave 1 interview was completed. The age classification date for a Wave 2 nonrespondent is two months after the last day of his/her anniversary month or the final date of the Wave 2 data collection (October 31, 2015), whichever is earlier.⁴ Age as of the age classification date was used to categorize Wave 1 youth who were nonrespondents in Wave 2. That is, a Wave 1 youth who did not respond in Wave 2 was classified as a continuing youth if his/her age was determined to be 17 or younger on the age classification date; otherwise he/she was classified as an aged-up adult. A similar classification rule was used for persons who were shadow youth at Wave 1. A Wave 1 shadow youth who did not respond in Wave 2 was classified as an aged-up youth if he/she was determined to have attained age 12 on or before the age classification date.

Table 2-1 displays the case counts according to their adult/youth/shadow youth classification in Wave 1 and their classification in Wave 2.

Table 2-1. Case counts by Wave 1 adult/youth/shadow youth classification and Wave 2 adult/youth/shadow youth classification

Wave 1 classification	Wave 2 classification	Unweighted count
Adult	Continuing adult	32,320
Youth	Aged-up adult	2,239
Youth	Continuing youth	11,412
Shadow youth	Aged-up youth	2,555
Shadow youth	Shadow youth	4,652

⁴ The purpose of the “anniversary month” concept is to provide a reasonable target period for completing the Wave 2 interview that would improve the likelihood of the interviews taking place one year after each individual’s Wave 1 interview, by giving the field interviewers a target for their data collection efforts. The Wave 2 anniversary month is defined as the calendar month containing the date of the earliest Wave 1 interview completed by a member of the study participant’s household, one year after the Wave 1 interview. Given the challenges of contacting and scheduling the interviews, the target period encompassed four months, starting with the month before the anniversary month and ending two months after the anniversary month; however, if necessary, efforts to complete the interview could continue past this period. In theory, data collection efforts for nonrespondents could continue up to the last day of the data collection period; because a reasonable cutoff date was needed for defining the ages of nonrespondents, the end of the target period was chosen as a consistent and justifiable standard.

The interview retention and recruitment rates are presented for demographic and tobacco use subgroups defined using Wave 1 characteristics. This is necessary because Wave 2 characteristics are not available for Wave 2 nonrespondents. All persons asked to provide a biospecimen at Wave 2 completed the Wave 2 Adult Interview so the subgroup definitions for the biospecimen response rates use Wave 2 data for characteristics where information was updated (age and tobacco use status), and Wave 1 data otherwise.

2.1.1 Retention Rates for Continuing Adults and Continuing Youth

This section reports retention rates for continuing adults, who completed the Adult Interview at Wave 1, and continuing youth, who completed the Youth Interview at Wave 1 and remained eligible for the Youth Interview at Wave 2.

Method

Consistent with the response rate calculation guidelines specified by the Office of Management and Budget in its “Standards and Guidelines for Statistical Surveys” (2006), the retention rate for Wave 2 was calculated as the ratio of the number of Wave 2 complete cases (or sufficient partials) to the number of cases eligible for the Wave 2 interview, which is essentially the American Association for Public Opinion Research (AAPOR) RR3 (AAPOR, 2015) response rate.

The retention rate for continuing adults, denoted as RR_{CA} , was calculated using the equations below:

$$RR_{CA} = C_{CA} / (C_{CA} + N_{CA} + e_{CA} * U_{CA})$$

$$e_{CA} = C_{CA} / (C_{CA} + I_{CA})$$

where

- C_{CA} = number of Wave 2 completed cases or sufficient partials among continuing adults;
- N_{CA} = number of Wave 2 nonrespondents known to be eligible among continuing adults;
- U_{CA} = number of Wave 2 nonrespondents with eligibility unknown among continuing adults;
- I_{CA} = number of Wave 2 ineligible cases among continuing adults; and

e_{CA} = estimated proportion of nonrespondents with unknown eligibility who were eligible among continuing adults.

The retention rate for continuing youth, denoted as RR_{CY} , was calculated using the equations below:

$$RR_{CY} = C_{CY} / (C_{CY} + N_{CY} + e_{CY} * U_{CY})$$

$$e_{CY} = C_{CY} / (C_{CY} + I_{CY})$$

where

- C_{CY} = number of Wave 2 completed cases or sufficient partials among continuing youth;
- N_{CY} = number of Wave 2 nonrespondents known to be eligible among Wave 1 youth respondents who were age 17 or younger on the age classification date;
- U_{CY} = number of Wave 2 nonrespondents with eligibility unknown among Wave 1 youth respondents who were age 17 or younger on the age classification date;
- I_{CY} = number of Wave 2 ineligible cases among Wave 1 youth respondents who were age 17 or younger on the age classification date; and
- e_{CY} = estimated proportion of nonrespondents with unknown eligibility who were eligible among Wave 1 youth respondents who were age 17 or younger on the age classification date.

Both unweighted and weighted retention rates were calculated. For the unweighted retention rates, the numbers of cases used in the calculations are the actual case counts. For the weighted retention rates, the numbers of cases used in the calculations are the sums of Wave 1 IPS weights, in accordance with AAPOR guidance (see AAPOR, 2015, p.51). The unweighted retention rate measures the success of field operations in obtaining responses from the sample group. The weighted retention rate estimates the proportion of the population represented by the sample group that would have responded if they all had been asked to participate in the study, and provides a measure of the potential impact of nonresponse on the quality of weighted estimates.

Results

Tables 2-2 and 2-3 provide unweighted and weighted retention rates by Wave 1 characteristics for continuing adults and continuing youth. In addition to the overall row, each table includes rows on age, sex, race, and ethnicity subgroups based on self-reported data from the Wave 1 Extended Interviews. Both tables also include rows on tobacco use status. Persons with missing values for these characteristics from Wave 1 were excluded from the response rate calculation for the particular characteristic.

Table 2-2. PATH Study Wave 2 retention rates by Wave 1 characteristics: Adult Interview (continuing adults)

Characteristic at Wave 1 ^a	C _{CA} : Adult Interview, completed (n)	I _{CA} : Ineligible (n)	N _{CA} : Nonresponse known to be eligible (n)	U _{CA} : Nonresponse with unknown eligibility (n)	Unweighted RR _{CA} : Unweighted retention rate for Wave 2 (%)	Weighted RR _{CA} : Weighted retention rate for Wave 2 (%)
Overall	26,459	263	4,047	1,551	82.6	83.1
Tobacco use status ^b						
Current established user	11,693	135	1,753	788	82.2	82.1
Not current established user	14,039	112	2,136	728	83.1	83.8
Age						
18-24	7,330	40	1,097	645	80.8	81.2
25-44	9,233	53	1,351	632	82.3	82.6
45-64	7,375	87	1,110	246	84.5	84.6
65+	2,516	83	483	28	83.1	82.9
Sex						
Male	13,080	162	2,167	900	81.1	81.8
Female	13,356	101	1,875	650	84.1	84.2
Race						
White alone	18,997	191	3,086	970	82.4	83.3
Black alone or in combination with some other race	4,647	37	516	338	84.5	84.4
Other	2,170	22	341	193	80.3	80.2
Ethnicity						
Hispanic	4,528	43	580	385	82.5	82.7
Non-Hispanic	21,539	213	3,391	1,145	82.6	83.3

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 2-3. PATH Study Wave 2 retention rates by Wave 1 characteristics: Youth Interview (continuing youth)

Characteristic at Wave 1 ^a	C _{cy} : Youth Interview, completed (n)	I _{cy} : Ineligible (n)	N _{cy} : Nonresponse known to be eligible (n)	U _{cy} : Nonresponse with unknown eligibility (n)	Unweighted RR _{cy} : Unweighted retention rate for Wave 2 (%)	Weighted RR _{cy} : Weighted retention rate for Wave 2 (%)
Overall	10,081	15	964	352	88.5	88.4
Tobacco use status ^b						
Ever user	1,665	4	170	91	86.5	86.6
Never user	8,006	11	737	245	89.1	89.0
Age						
12-13	4,164	6	381	133	89.0	88.9
14-17	5,917	9	582	219	88.1	88.0
Sex						
Male	5,164	11	510	155	88.6	88.5
Female	4,892	4	445	196	88.4	88.4
Race						
White alone	6,497	7	662	190	88.4	88.4
Black alone or in combination with some other race	1,882	2	149	107	88.0	87.7
Other	1,143	5	95	36	89.7	89.4
Ethnicity						
Hispanic	2,907	7	244	131	88.6	88.5
Non-Hispanic	6,982	7	698	213	88.5	88.4

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

The weighted retention rates are approximately 83 percent for continuing adults and 88 percent for continuing youth. The unweighted retention rates are approximately the same as the weighted retention rates. These rates are moderately lower than the projected retention rates reported in the Revision Request to OMB for Wave 2 (i.e., 86 percent for continuing adults and 90 percent for continuing youth). For continuing adults, it appears that females had higher retention rates than males, 18-24 years olds had lower retention rates than older age groups, adults who identified as Black or White alone had higher retention rates than those of “other” race, and current established tobacco users at Wave 1 had slightly lower retention rates than those who were not. The retention rates are similar for most the subgroups of continuing youth, but are slightly higher among Wave 1 never users of tobacco. Slight variation in response rates by subgroups is to be expected in large-scale data collection efforts. Although there is some variability among the retention rates for the various subgroups of interest, none of the subgroup differences in Tables 2-2 and 2-3 suggests a reason for concern.

2.1.2 Recruitment Rates for Aged-up Adults and Aged-up Youth

This section reports recruitment rates for aged-up adults, who completed the Youth Interview at Wave 1 and were eligible for the Adult Interview at Wave 2, and aged-up youth, who were shadow youth at Wave 1 and were eligible for the Youth Interview at Wave 2. The Wave 2 Youth Interview was the first interview for responding aged-up youth, and aged-up adult respondents completed the Adult Interview for the first time.

Method

The methods described in Section 2.1.1 for calculating the retention rates were also used to calculate the recruitment rates for aged-up adults and aged-up youth. For aged-up adults, the recruitment rate, denoted as RR_{AUA} , was calculated using the equations below:

$$RR_{AUA} = C_{AUA} / (C_{AUA} + N_{AUA} + e_{AUA} * U_{AUA})$$

$$e_{AUA} = C_{AUA} / (C_{AUA} + I_{AUA})$$

where

- C_{AUA} = number of Wave 2 completed cases or sufficient partials among persons who completed the Youth Interview at Wave 1 and were administered the Adult Interview at Wave 2;
- N_{AUA} = number of Wave 2 nonrespondents known to be eligible among Wave 1 youth respondents who attained age 18 by the age classification date;
- U_{AUA} = number of Wave 2 nonrespondents with eligibility unknown among Wave 1 youth respondents who attained age 18 by the age classification date;
- I_{AUA} = number of Wave 2 ineligible cases among Wave 1 youth respondents who attained age 18 by the age classification date; and
- e_{AUA} = estimated proportion of nonrespondents with unknown eligibility who were eligible among Wave 1 youth respondents who attained age 18 by the age classification date.

The recruitment rate for aged-up youth, denoted as RR_{AUY} , was calculated using the equations below:

$$RR_{AUY} = C_{AUY} / (C_{AUY} + N_{AUY} + e_{AUY} * U_{AUY})$$

$$e_{AUY} = C_{AUY} / (C_{AUY} + I_{AUY})$$

where

- C_{AUY} = number of Wave 2 completed cases or sufficient partials among persons who were shadow youth at Wave 1 and were administered the Youth Interview at Wave 2;
- N_{AUY} = number of Wave 2 nonrespondents known to be eligible among Wave 1 shadow youth who were age 12 by the age classification date;
- U_{AUY} = number of Wave 2 nonrespondents with eligibility unknown among Wave 1 shadow youth who were age 12 by the age classification date;
- I_{AUY} = number of Wave 2 ineligible cases among Wave 1 shadow youth who were age 12 by the age classification date; and
- e_{AUY} = estimated proportion of nonrespondents with unknown eligibility who were eligible among Wave 1 shadow youth who were age 12 by the age classification date.

As for the retention rates, both unweighted and weighted recruitment rates were calculated. For the unweighted recruitment rates, the numbers of cases used in the calculations are the actual case counts. For the weighted recruitment rates, the numbers of cases used in the calculations are the sums of Wave 1 IPS weights (AAPOR, 2015).

Results

Table 2-4 provides recruitment rates by Wave 1 characteristics for the Adult Interview for aged-up adults, and Table 2-5 provides recruitment rates by Wave 1 characteristics for the Youth Interview for aged-up youth. In addition to the overall row, each table includes rows on sex, race, and ethnicity subgroups; Table 2-4 also includes rows on tobacco use status (which is not available for aged-up youth). There are no rows corresponding to age subgroups in Table 2-4 or Table 2-5, because most of the aged-up adults are 18 years old and most of the aged-up youth are 12 years old. Information from the Wave 1 Extended Youth Interview was used to define the demographic and tobacco-use characteristics for the aged-up adults, and information from the Wave 1 Household Screener was used to define the demographic characteristics for the aged-up youth. Persons with missing values for these characteristics were excluded from the response rate calculation for that characteristic.

The weighted recruitment rate is approximately 86 percent for aged-up adults, which is slightly higher than the projected recruitment rate for this group (85 percent) provided in the Revision Request to OMB for Wave 2. The weighted recruitment rate is approximately 82 percent for aged-up youth, which is lower than the projected recruitment for this group (88 percent) in the Wave 2 Revision Request. For both aged-up adults and aged-up youth, the variability of recruitment rates by subgroups is small. For aged-up adults, the differences in recruitment rates by subgroup range from 0.8 percentage points for tobacco use status to 2.3 percentage points for race/ethnicity. For aged-up adults, the differences in recruitment rates by subgroup are 0.8 percentage points for sex and 1.1 percentage points for race/ethnicity.

Table 2-4. PATH Study Wave 2 recruitment rates by Wave 1 characteristics: Adult Interview (aged-up adults)

Characteristic at Wave 1 ^a	C _{AUA} : Adult Interview, completed (n)	I _{AUA} : Ineligible (n)	N _{AUA} : Nonresponse known to be eligible (n)	U _{AUA} : Nonresponse with unknown eligibility (n)	Unweighted RR _{AUA} : Unweighted recruitment rate for Wave 2 (%)	Weighted RR _{AUA} : Weighted recruitment rate for Wave 2 (%)
Overall	1,916	8	219	96	85.9	85.7
Tobacco use status ^b						
Ever user	811	2	80	46	86.6	86.2
Never user	1,061	6	133	47	85.5	85.4
Sex						
Male	957	5	123	46	85.0	84.7
Female	956	2	96	50	86.8	86.6
Race/ethnicity						
Non-Hispanic White alone	954	3	136	37	84.7	84.5
Other	945	5	83	58	87.0	86.8

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-5. PATH Study Wave 2 recruitment rates by Wave 1 characteristics: Youth Interview (aged-up youth)

Characteristic at Wave 1 ^a	C _{AUY} : Youth Interview, completed (n)	I _{AUY} : Ineligible (n)	N _{AUY} : Nonresponse known to be eligible (n)	U _{AUY} : Nonresponse with unknown eligibility (n)	Unweighted RR _{AUY} : Unweighted recruitment rate for Wave 2 (%)	Weighted RR _{AUY} : Weighted recruitment rate for Wave 2 (%)
Overall	2,091	5	367	92	82.0	82.1
Sex						
Male	1,055	3	179	58	81.7	81.7
Female	1,036	2	188	34	82.4	82.5
Race/ethnicity						
Non-Hispanic White alone	1,008	2	190	25	82.4	82.6
Other	1,078	3	176	67	81.6	81.5

^a The characteristics are as reported in the Household Screener at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

2.1.3 Biospecimen Collections

This section addresses the response rates for the collection of urine and blood specimens from continuing adults and aged-up adults who completed a Wave 2 Adult Interview and were asked to provide a specimen. Biospecimens are intended to provide a basis for the assessment of between-person differences and within-person changes in markers of tobacco exposure, and to detect and compare indicators of conditions and related disease processes associated with the use of tobacco products. Field interviewers collected the urine specimens; on separate visits, phlebotomists collected the blood specimens. Among Wave 2 Adult Interview respondents, a subsample of continuing adults were asked to provide a urine specimen and all aged-up adults were asked to provide urine and blood specimens.

Method

Tables 2-6 and 2-7 provide overall unweighted response rates⁵ for the biospecimen collections, and response rates by tobacco use status and demographic subgroups, for continuing adults and aged-up adults, respectively. The response rates are conditional on a completed Wave 2 Adult Interview and the adult having been asked to provide a specimen. The response rates were calculated using the following formula:

$$\text{RRU} = \frac{\text{(Number of adults who provided a specimen)}}{\text{(Number of adults from whom a specimen was requested)}}$$

This is the same formula used to compute the biospecimen response rates for Wave 1. However, all Adult Interview respondents were asked to provide urine and blood biospecimens in Wave 1. The denominator for the Wave 2 urine response rate in Table 2-6 is the 12,569 continuing adults who completed the Wave 2 Adult Interview and were asked for a urine specimen. The denominator for the Wave 2 urine and blood response rates in Table 2-7 is the 1,916 aged-up adults who completed the Wave 2 Adult Interview.

⁵ Weighted response rates are not provided because the subset of continuing adults asked to provide a urine specimen at Wave 2 does not represent a readily interpretable portion of the population. Similarly, while most aged-up adults were age 18 at Wave 2, some were older due to the actual time elapsed between the Wave 1 and Wave 2 interviews.

Table 2-6. PATH Study Wave 2 response rates by respondent characteristics: Urine collection (continuing adults)

Characteristic ^a	A: Adults requested to provide urine (n)	Urine	
		B: Collected (n)	Unweighted response rate for Wave 2 ^c (%)
Overall	12,569	12,113	96.4
Tobacco use status ^b			
Current established user	7,977	7,725	96.8
Not current established user	4,533	4,332	95.6
Age			
18-24	3,326	3,195	96.1
25-44	4,948	4,789	96.8
45-64	3,450	3,325	96.4
65+	845	804	95.1
Sex			
Male	6,491	6,229	96.0
Female	6,070	5,878	96.8
Race			
White alone	8,855	8,540	96.4
Black alone or in combination	2,400	2,316	96.5
Other	1,032	989	95.8
Ethnicity			
Hispanic	2,125	2,025	95.3
Non-Hispanic	10,272	9,927	96.6

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Adult Extended Interview. The age information and tobacco use status are as reported in the Wave 2 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 2 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

^c Response rate = B/A.

Table 2-7. PATH Study Wave 2 response rates by respondent characteristics: Biospecimen collections (aged-up adults)

Characteristic ^a	A: Adult Interviews completed (n)	Urine		Blood	
		B: Collected (n)	Unweighted response rate for Wave 2 ^c (%)	B: Collected (n)	Unweighted response rate for Wave 2 ^c (%)
Overall	1,916	1,587	82.8	908	47.4
Tobacco use status ^b					
Current established user	328	295	89.9	177	54.0
Not current established user	1,578	1,285	81.4	727	46.1
Sex					
Male	957	792	82.8	457	47.8
Female	956	792	82.8	450	47.1
Race					
White alone	1,257	1,034	82.3	596	47.4
Black alone or in combination	348	299	85.9	161	46.3
Other	212	173	81.6	109	51.4
Ethnicity					
Hispanic	515	426	82.7	251	48.7
Non-Hispanic	1,384	1,145	82.7	648	46.8

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The tobacco use status is as reported in the Wave 2 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 2 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A ‘current established user’ of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

^c Response rate = B/A.

In addition to the overall row, each table includes rows on tobacco use status, sex, race, and ethnicity subgroups. Table 2-6 includes rows on age subgroups; this is not necessary for Table 2-7 because the age range among Wave 2 aged-up adults is narrow. Information from the Wave 2 Adult Interview was used to define the age categories and tobacco use status; information from the Wave 1 Adult Interview was used to define the sex, race, and ethnicity categories (because these characteristics were not re-asked of Wave 2 adults). Adults with missing values for such characteristics were excluded from the response rate calculation for that characteristic.

Results

As shown in Table 2-6, the unweighted response rate for **urine** was approximately 96 percent among continuing adults from whom a specimen was requested. This is considerably higher than the

80 percent response rate projected in the Revision Request to OMB for Wave 2. The response rates vary little across the various tobacco use status and demographic subgroups.

As shown in Table 2-7, the unweighted response rate for **urine** among aged-up adults was approximately 83 percent. This is considerably higher than the projected response rate of 69 percent. It appears that current established tobacco users at Wave 2 had a higher tendency to provide urine specimens than those who were not, and Black aged-up adults had a higher response rate for the urine specimen collection than aged-up adults of other races. There is little variation in urine specimen response rates by sex or ethnicity among aged-up adults.

The unweighted response rate for **blood** among aged-up adults was approximately 47 percent. This is higher than the projected response rate of 45 percent. The differential unweighted response rate for subgroups of respondents ranges from 1 percentage point for sex to 8 percentage points for tobacco use status.

2.2 Nonresponse Bias Analysis

The standard approach for an analysis of nonresponse bias in the first follow-up of a longitudinal cohort study such as the PATH Study is to compare Wave 2 respondents with Wave 2 nonrespondents with respect to characteristics from Wave 1 (Bose and West, 2002; Javitz and Wagner, 2005; Brownstein et al., 2009). By so doing, the study can assess the extent to which differential nonresponse among population subgroups may affect estimates. Results are presented on the characteristics of respondents to the Wave 2 Adult and Youth Interviews, and on aged-up adults from whom biospecimens were collected at Wave 2. The response rates for urine collection from continuing adults are above 95 percent and vary little for all the subgroups in Table 2-6; further analysis of potential nonresponse bias for this aspect of Wave 2 biospecimen collection is therefore not warranted.

Analyses of bias for the Wave 2 interviews are presented for demographic and tobacco use subgroups based on their Wave 1 characteristics because Wave 2 characteristics are not available for Wave 2 nonrespondents. All persons asked to provide a biospecimen at Wave 2 completed the Wave 2 Adult Interview so the subgroup definitions for the biospecimen nonresponse analysis use Wave 2 data for characteristics where information was updated (tobacco use status), and Wave 1 data otherwise.

The PATH Study measures a range of tobacco use behaviors; many of these variables are not available in other studies. However, responses to the PATH Study questions on current cigarette smoking can be compared with estimates from other surveys that ask about cigarette smoking behavior. A separate component of the nonresponse bias analysis investigates possible differences between cigarette smoking estimates calculated from Wave 2 of the PATH Study and independent estimates of those quantities from other studies for the most similar time periods for which estimates were available.

Both components of the analysis of nonresponse bias are based on the full set of Wave 2 data.

2.2.1 Method

The method used in the PATH Study to assess potential interview nonresponse bias begins by comparing estimates of Wave 1 demographic characteristics and tobacco use for Wave 2 respondents with corresponding estimates for Wave 2 nonrespondents.⁶ These comparisons are made separately for continuing adults, continuing youth, aged-up adults, and aged-up youth. To assess potential nonresponse bias associated with the collection of urine and blood specimens from aged-up adults at Wave 2, estimates of Wave 1 demographic characteristics and Wave 2 tobacco use are compared for specimen providers and non-providers, separately for each type of biospecimen.

Sections 2.2 and 2.3 of the 2015 Interim Report describe the weight construction for Wave 1 of the PATH Study. The final raked weights from Wave 1 were designed to reduce the potential nonresponse bias from Wave 1. For Wave 2, the interview nonresponse bias analysis uses the raked weights from Wave 1. Differences between the weighted estimates of Wave 1 characteristics for Wave 2 respondents and nonrespondents therefore identify characteristics that might be associated with nonresponse bias due to attrition between the first two waves of the study, after compensating for Wave 1 nonresponse and possible undercoverage.

Section 2.3.1 of this Interim Report describes the weight construction for Wave 2 of the PATH Study. The final raked weights from Wave 2 were designed to reduce the potential nonresponse bias from Wave 2. For Wave 2, the biospecimen nonresponse bias analysis uses the raked weights from Wave 2 because all aged-up adults asked to provide urine and blood specimens were Wave 2 Adult

⁶ Persons determined to be ineligible for Wave 2 (see Section 1.2.2) are not included in this analysis.

Interview respondents. Differences between the weighted estimates of characteristics for Wave 2 specimen providers and non-providers therefore identify characteristics that might be associated with bias due to biospecimen nonresponse at Wave 2, after accounting for Wave 1 nonresponse and possible undercoverage, and attrition between the first two waves of the study.

Wave 2 cigarette smoking estimates were compared to estimates based on data from the following surveys: TUS-CPS, 2010-2011; NHIS, 2014; NHANES, 2013-2014; NSDUH, 2014; and NYTS, 2014. Appendix A describes the questions used to define current cigarette smoking on each of these surveys as well as the PATH Study, and outlines differences in target populations among these surveys and the PATH Study. This second component of the nonresponse bias analysis uses the Wave 1 IPS weights. Differences between the weighted estimates of current cigarette smoking behavior reported by Wave 2 respondents therefore identify characteristics that might be associated with bias due to nonresponse at Wave 1 or attrition between the first two waves of the study.

In the tables presented in Section 2.2.2, the unweighted counts and estimates of percentages calculated using weights exclude respondents with missing values for that item. The estimates calculated from other surveys that are used for comparison purposes also exclude missing values. The proportions of item missingness are generally very low in both the PATH Study and the surveys that are used for comparison purposes.

Point estimates for the PATH Study were calculated using the Wave 1 final weights or Wave 1 IPS weights or Wave 2 final weights, as described above. The corresponding replicate weights were used to calculate variances, and account for the complex sampling features of stratification and clustering. Precisions for the PATH Study estimates are reported using 95 percent confidence intervals based on the modified Wilson confidence interval approach (Wilson, 1927; SAS Institute, 2013). Estimates from TUS-CPS, NHANES, NHIS, NSDUH, and NYTS⁷ also have sampling error, so 95 percent confidence intervals are reported for the estimates from those surveys as well.⁸ SAS software version 9.4 was used to calculate all point estimates and confidence intervals.

⁷ Estimates from TUS-CPS were obtained from United States Department of Commerce (2012) published tables. Estimates from NHANES, NHIS, NYTS, and NSDUH were calculated from their respective public use files (Centers for Disease Control and Prevention (2015a, 2015b, 2015c), Center for Behavioral Health Statistics and Quality (2015a)).

⁸ The modified Wilson confidence interval approach was used to compute the 95 percent confidence intervals for NHANES, NHIS, NSDUH, and NYTS; the TUS-CPS documentation does not state what method was used to produce the confidence intervals that appear in the published tables. If the confidence interval from the PATH Study estimate does not overlap with the confidence interval from the comparison study, then the results are significantly different at the 0.05 significance level. Schenker and Gentleman (2001) show that this results in a conservative test. In

2.2.2 Results

Tables 2-8 and 2-9 compare Wave 1 demographic characteristics⁹ and tobacco use rates, respectively, for Wave 2 continuing adult respondents with nonrespondents. Tables 2-10 and 2-11 present similar comparisons for continuing youth at Wave 2. Tables 2-12 and 2-13 present the comparisons for Wave 2 aged-up adult respondents and nonrespondents. Table 2-14 compares Wave 1 demographic characteristics for Wave 2 aged-up youth respondents with nonrespondents (tobacco use was not measured for shadow youth at Wave 1). If nonresponse at Wave 2 is not associated with demographic characteristics or Wave 1 tobacco use, then the estimates calculated using the Wave 1 final weights would be expected to be similar for Wave 2 respondents and nonrespondents.

The results in Table 2-8 show that males, 18-24 year-olds, and those with high school education are underrepresented among continuing adult respondents; and 45-64 year-olds, persons with health insurance, and those with at least a bachelor's degree are overrepresented among respondents. Table 2-9 shows that estimates of Wave 1 current established tobacco use are lower overall, and for males, 18-44 year-olds, and non-Hispanic Whites among continuing adult respondents. Tables 2-10 and 2-11 show that estimates of Wave 1 ever use of tobacco are lower for continuing youth respondents than nonrespondents. Table 2-11 shows that estimates of Wave 1 ever use of tobacco are also lower for females, 14-17 year-olds, and non-Hispanic Whites among continuing youth respondents. If no further weighting adjustments were performed then to the extent that these characteristics are associated with the PATH Study's outcomes, those outcomes may be affected by nonresponse bias. However, this concern is addressed by the weighting adjustments and results described in Section 2.3.

No evidence of potential nonresponse bias was found for aged-up adults or aged-up youth based on the estimates in Tables 2-12, 2-13, and 2-14 because all the confidence intervals for estimates of difference between Wave 2 respondents and nonrespondents include zero.

general, if a 95 percent confidence interval for the percentage of adults who are current cigarette smokers from the PATH Study includes a fixed value X, then a hypothesis test of the null hypothesis that the percentage of adults who are current smokers equals X would have p-value > 0.05 and therefore the difference between the PATH Study estimate and the estimate from the external survey is not statistically significant. No adjustments were made for multiple testing; however, results are reported here for all of the comparisons performed, and all of these were pre-planned comparisons.

⁹ Health insurance coverage is also included in Table 2-8.

Table 2-8. Comparison of Wave 2 continuing Adult Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 2 continuing adult nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Sex					
Male	13,080	47.2% [46.6%, 47.8%]	3,067	51.4% [49.6%, 53.2%]	-4.2% [-6.4%, -2.0%]
Female	13,356	52.8% [52.2%, 53.4%]	2,525	48.6% [46.8%, 50.4%]	4.2% [2.0%, 6.4%]
Age group					
18-24	7,330	12.8% [12.4%, 13.2%]	1,742	14.2% [13.3%, 15.1%]	-1.4% [-2.4%, -0.3%]
25-44	9,233	34.2% [33.6%, 34.7%]	1,983	35.5% [34.2%, 36.9%]	-1.3% [-3.0%, 0.3%]
45-64	7,375	35.1% [34.6%, 35.7%]	1,356	31.9% [30.4%, 33.5%]	3.2% [1.3%, 5.1%]
65+	2,516	17.9% [17.4%, 18.4%]	511	18.4% [16.9%, 20.0%]	-0.5% [-2.4%, 1.4%]
Race/ethnicity					
Non-Hispanic White alone	15,767	66.2% [65.6%, 66.7%]	3,371	65.1% [63.5%, 66.6%]	1.1% [-0.8%, 3.0%]
Other	10,252	33.8% [33.3%, 34.4%]	2,116	34.9% [33.4%, 36.5%]	-1.1% [-3.0%, 0.8%]
Health insurance					
Yes	21,269	85.9% [85.3%, 86.5%]	4,281	83.3% [81.8%, 84.6%]	2.6% [1.1%, 4.2%]
No	4,924	14.1% [13.5%, 14.7%]	1,186	16.7% [15.4%, 18.2%]	-2.6% [-4.2%, -1.1%]
Education					
< HS or GED	5,279	16.4% [15.9%, 16.8%]	1,097	17.8% [16.5%, 19.1%]	-1.4% [-2.9%, 0.1%]
HS	6,056	23.8% [23.3%, 24.3%]	1,421	26.5% [24.9%, 28.1%]	-2.7% [-4.5%, -0.8%]
Some college, no degree	9,308	31.2% [30.6%, 31.7%]	1,923	30.7% [29.2%, 32.2%]	0.5% [-1.3%, 2.2%]

Table 2-8. Comparison of Wave 2 continuing Adult Interview respondents with nonrespondents (continued)

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 2 continuing adult nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Education (continued)					
Bachelor degree +	5,684	28.6% [28.1%, 29.2%]	1,087	25.0% [23.4%, 26.8%]	3.6% [1.6%, 5.6%]
Tobacco use status ^b					
Current established user	11,693	23.5% [22.9%, 24.2%]	2,541	25.6% [24.5%, 26.8%]	-2.1% [-3.4%, -0.8%]
Not current established user	14,039	76.5% [75.8%, 77.1%]	2,864	74.4% [73.2%, 75.5%]	2.1% [0.8%, 3.4%]

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 2-9. Comparison of Wave 1 tobacco use* rates for Wave 2 continuing Adult Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 2 continuing adult nonrespondents		Difference in weighted percentages [respondents - nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Overall	25,732	23.5% [22.9%, 24.2%]	5,405	25.6% [24.5%, 26.8%]	-2.1% [-3.4%, -0.8%]
Sex					
Male	12,738	29.0% [28.0%, 30.0%]	2,961	33.2% [31.5%, 34.9%]	-4.1% [-6.1%, -2.1%]
Female	12,973	18.6% [17.8%, 19.3%]	2,439	17.7% [16.2%, 19.2%]	0.9% [-0.6%, 2.4%]
Age group					
18-24	7,214	28.2% [26.7%, 29.7%]	1,708	32.6% [30.2%, 35.1%]	-4.4% [-6.9%, -2.0%]
25-44	9,052	28.6% [27.5%, 29.6%]	1,938	31.1% [29.1%, 33.2%]	-2.5% [-4.9%, -0.1%]
45-64	7,113	23.0% [22.0%, 24.0%]	1,293	24.5% [22.2%, 26.9%]	-1.5% [-3.7%, 0.7%]
65+	2,348	10.8% [9.6%, 12.1%]	460	10.4% [7.9%, 13.5%]	0.4% [-1.9%, 2.7%]
Race/ethnicity					
Non-Hispanic White alone	15,428	24.6% [23.7%, 25.6%]	3,272	28.7% [27.1%, 30.2%]	-4.1% [-5.6%, -2.5%]
Other	9,923	21.3% [20.4%, 22.2%]	2,039	20.3% [18.4%, 22.4%]	1.0% [-1.3%, 3.3%]

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

* Defined as current established use of tobacco. A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 2-10. Comparison of Wave 2 continuing Youth Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 continuing youth respondents		Wave 2 continuing youth nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	5,164	51.4% [50.5%, 52.4%]	665	50.6% [47.9%, 53.3%]	0.8% [-2.3%, 3.9%]
Female	4,892	48.6% [47.6%, 49.5%]	641	49.4% [46.7%, 52.1%]	-0.8% [-3.9%, 2.3%]
Age group					
12-13	4,164	41.1% [40.1%, 42.1%]	514	39.4% [36.8%, 42.1%]	1.7% [-1.0%, 4.3%]
14-17	5,917	58.9% [57.9%, 59.9%]	801	60.6% [57.9%, 63.2%]	-1.7% [-4.3%, 1.0%]
Race/ethnicity					
Non-Hispanic White alone	4,731	54.6% [53.6%, 55.5%]	616	53.8% [50.7%, 57.0%]	0.7% [-2.9%, 4.3%]
Other	5,141	45.4% [44.5%, 46.4%]	667	46.2% [43.0%, 49.3%]	-0.7% [-4.3%, 2.9%]
Tobacco use status^b					
Ever user	1,665	16.9% [16.0%, 17.9%]	261	20.4% [18.2%, 22.7%]	-3.4% [-5.7%, -1.2%]
Never user	8,006	83.1% [82.1%, 84.0%]	982	79.6% [77.3%, 81.8%]	3.4% [1.2%, 5.7%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-11. Comparison of Wave 1 tobacco use* rates for Wave 2 continuing Youth Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 continuing youth respondents		Wave 2 continuing youth nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Overall	9,671	16.9% [16.0%, 17.9%]	1,243	20.4% [18.2%, 22.7%]	-3.4% [-5.7%, -1.2%]
Sex					
Male	4,922	18.0% [16.9%, 19.2%]	619	21.4% [18.2%, 25.0%]	-3.4% [-7.0%, 0.2%]
Female	4,727	15.8% [14.7%, 17.0%]	616	19.4% [16.5%, 22.8%]	-3.7% [-6.9%, -0.4%]
Age group					
12-13	3,923	7.8% [6.9%, 9.0%]	473	7.9% [5.7%, 10.7%]	-0.0% [-2.7%, 2.7%]
14-17	5,748	23.1% [21.9%, 24.3%]	769	28.1% [24.9%, 31.6%]	-5.1% [-8.4%, -1.8%]
Race/ethnicity					
Non-Hispanic White alone	4,555	17.6% [16.4%, 19.0%]	583	24.0% [20.6%, 27.8%]	-6.4% [-9.8%, -2.9%]
Other	4,930	16.3% [15.1%, 17.4%]	633	16.6% [13.7%, 20.1%]	-0.4% [-3.7%, 3.0%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

* Defined as ever use of tobacco. An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-12. Comparison of Wave 2 aged-up Adult Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 aged-up adult respondents		Wave 2 aged-up adult nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	957	50.4% [48.1%, 52.6%]	169	54.3% [48.8%, 59.8%]	-4.0% [-10.3%, 2.3%]
Female	956	49.6% [47.4%, 51.9%]	146	45.7% [40.2%, 51.2%]	4.0% [-2.3%, 10.3%]
Race/ethnicity					
Non-Hispanic White alone	954	54.7% [52.4%, 56.9%]	173	58.9% [52.5%, 65.0%]	-4.2% [-11.4%, 3.0%]
Other	945	45.3% [43.1%, 47.6%]	141	41.1% [35.0%, 47.5%]	4.2% [-3.0%, 11.4%]
Tobacco use status^b					
Ever user	811	42.9% [40.4%, 45.5%]	126	40.5% [34.7%, 46.5%]	2.4% [-4.3%, 9.2%]
Never user	1,061	57.1% [54.5%, 59.6%]	180	59.5% [53.5%, 65.3%]	-2.4% [-9.2%, 4.3%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-13. Comparison of Wave 1 tobacco use* rates for Wave 2 aged-up Adult Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 aged-up adult respondents		Wave 2 aged-up adult nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Overall	1,872	42.9% [40.4%, 45.5%]	306	40.5% [34.7%, 46.5%]	2.4% [-4.3%, 9.2%]
Sex					
Male	933	46.6% [42.8%, 50.3%]	164	44.0% [35.7%, 52.5%]	2.6% [-7.4%, 12.6%]
Female	936	39.4% [36.3%, 42.5%]	142	36.4% [27.9%, 45.9%]	3.0% [-6.5%, 12.5%]
Race/ethnicity					
Non-Hispanic White alone	936	46.7% [42.9%, 50.5%]	169	43.8% [36.5%, 51.3%]	2.9% [-4.8%, 10.6%]
Other	919	38.1% [34.2%, 42.3%]	136	36.0% [26.8%, 46.4%]	2.2% [-9.4%, 13.7%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

* Defined as ever use of tobacco. An ‘ever user’ is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-14. Comparison of Wave 2 aged-up Youth Interview respondents with nonrespondents

Characteristic at Wave 1 ^a	Wave 2 aged-up youth respondents		Wave 2 aged-up youth nonrespondents		Difference in weighted percentages [respondents – nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using shadow youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using shadow youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	1,055	49.9% [47.8%, 52.1%]	237	51.4% [46.8%, 55.9%]	-1.4% [-6.9%, 4.0%]
Female	1,036	50.1% [47.9%, 52.2%]	222	48.6% [44.1%, 53.2%]	1.4% [-4.0%, 6.9%]
Race/ethnicity					
Non-Hispanic White alone	1,008	52.4% [50.2%, 54.5%]	215	51.3% [46.7%, 55.9%]	1.1% [-4.4%, 6.5%]
Other	1,078	47.6% [45.5%, 49.8%]	243	48.7% [44.1%, 53.3%]	-1.1% [-6.5%, 4.4%]

^a The characteristics are as reported in the Household Screener at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

Table 2-15 compares the characteristics of aged-up adults who provided a urine specimen with those of aged-up adults who did not provide a urine specimen. Table 2-16 presents a similar analysis for the collection of blood specimens from aged-up adults. If nonresponse to the Wave 2 biospecimen collection is not associated with demographic characteristics or Wave 2 tobacco use, then the estimates calculated using the Wave 2 final weights would be expected to be similar for Wave 2 specimen providers and non-providers. Based on the results in Tables 2-15 and 2-16, no evidence of nonresponse bias was found for sex and race/ethnicity. However, for both urine and blood, current established tobacco users are overrepresented among aged-up adults who provided specimens at Wave 2. This finding is consistent with the biospecimen response rates shown in Table 2-7.

Tables 2-17 and 2-18 address the second component of the nonresponse bias analysis and compare estimates of cigarette smoking calculated from Wave 2 of the PATH Study to independent estimates of those quantities from other studies.

Table 2-17 presents the estimates of prevalence of current cigarette smoking¹⁰ for adults based on the Wave 2 Adult Interview, for the adult population as a whole and for subgroups. These estimates are accompanied by 95 percent confidence intervals for the percentage of current cigarette smokers for the PATH Study estimates. The unweighted estimates are much higher than the weighted estimates due to the oversampling of adult tobacco users at Wave 1. The last five columns are the estimates of smoking prevalence from TUS-CPS, NHIS, NHANES, and NSDUH, respectively, along with 95 percent confidence intervals from those surveys. Note that these estimates exclude missing values.

The estimates of current smoking prevalence differ from survey to survey. Many potential reasons can explain these disparities, including that each survey has sampling error. Beyond that, however, the surveys differ in question order, context, design, mode of administration, and year of most recent data collection.

In general, the TUS-CPS estimates of smoking prevalence are lower than estimates from the other surveys, including the PATH Study. This may be related to the proxy responses used in the TUS-CPS. The rotation group structure of the TUS-CPS may result in underestimates of smoking prevalence, as smokers are more likely to drop out over the course of the panel survey (Song, 2013).

¹⁰ For the PATH Study, following common practice for tobacco surveys, a current cigarette smoker is someone who (1) has smoked at least 100 cigarettes in his or her lifetime and (2) currently smokes cigarettes every day or some days. The questions used to define current cigarette smoking for each survey are provided in Appendix A.

Table 2-15. Comparison of Wave 2 aged-up adult urine specimen providers with non-providers of urine

Characteristic ^a	Wave 2 aged-up adult urine providers		Wave 2 aged-up adult non-providers of urine		Difference in weighted percentages [providers – non-providers] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	
Sex					
Male	792	50.8% [48.3%, 53.2%]	165	50.5% [45.0%, 56.0%]	0.3% [-6.3%, 6.9%]
Female	792	49.2% [46.8%, 51.7%]	164	49.5% [44.0%, 55.0%]	-0.3% [-6.9%, 6.3%]
Race/ethnicity					
Non-Hispanic White alone	781	54.8% [52.4%, 57.3%]	173	56.2% [50.8%, 61.5%]	-1.4% [-7.3%, 4.5%]
Other	790	45.2% [42.7%, 47.6%]	155	43.8% [38.5%, 49.2%]	1.4% [-4.5%, 7.3%]
Tobacco use status^b					
Current established user	295	18.7% [16.5%, 21.0%]	33	10.7% [6.9%, 16.1%]	8.0% [2.7%, 13.2%]
Not current established user	1,285	81.3% [79.0%, 83.5%]	293	89.3% [83.9%, 93.1%]	-8.0% [-13.2%, -2.7%]

^a The sex and race/ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The tobacco use status is as reported in the Wave 2 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 2 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

Table 2-16. Comparison of Wave 2 aged-up adult blood specimen providers with non-providers of blood

Characteristic ^a	Wave 2 aged-up adult blood providers		Wave 2 aged-up adult non-providers of blood		Difference in weighted percentages [providers – non-providers] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	
Sex					
Male	457	52.0% [48.7%, 55.2%]	500	49.6% [46.6%, 52.7%]	2.4% [-2.7%, 7.4%]
Female	450	48.0% [44.8%, 51.3%]	506	50.4% [47.3%, 53.4%]	-2.4% [-7.4%, 2.7%]
Race/ethnicity					
Non-Hispanic White alone	445	55.1% [51.8%, 58.3%]	509	55.1% [52.0%, 58.2%]	-0.0% [-5.4%, 5.4%]
Other	454	44.9% [41.7%, 48.2%]	491	44.9% [41.8%, 48.0%]	0.0% [-5.4%, 5.4%]
Tobacco use status^b					
Current established user	177	20.2% [17.2%, 23.5%]	151	14.7% [12.5%, 17.3%]	5.4% [1.4%, 9.5%]
Not current established user	727	79.8% [76.5%, 82.8%]	851	85.3% [82.7%, 87.5%]	-5.4% [-9.5%, -1.4%]

^a The sex and race/ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The tobacco use status is as reported in the Wave 2 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 2 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

Table 2-17. Current cigarette smoking based on Wave 2 Adult Interview respondents: IPS weights

Characteristic at Wave 2	Unweighted count	PATH Study: Unweighted percentage	PATH Study: Weighted percentage, using adult IPS weights [95% confidence interval]	Percentage from 2010-2011 TUS-CPS [95% confidence interval]	Percentage from 2014 NHIS [95% confidence interval]	Percentage from 2013-2014 NHANES [95% confidence interval]	Percentage from 2014 NSDUH, original definition ^a [95% confidence interval]	Percentage from 2014 NSDUH, modified definition ^a [95% confidence interval]
Current smoker	28,337	34.2%	19.0% [18.1%, 19.9%]	16.1% [15.8%, 16.3%]	16.7% [16.1%, 17.4%]	20.0% [17.8%, 22.3%]	22.7% [22.1%, 23.2%]	20.9% [20.3%, 21.4%]
Current smoker, male	14,014	35.3%	21.2% [20.2%, 22.3%]	18.0% [17.7%, 18.4%]	18.8% [18.0%, 19.7%]	21.6% [19.4%, 24.1%]	25.5% [24.7%, 26.3%]	23.6% [22.8%, 24.4%]
Current smoker, female	14,297	33.3%	17.2% [16.2%, 18.1%]	14.2% [13.9%, 14.5%]	14.8% [13.9%, 15.7%]	18.4% [15.6%, 21.6%]	20.1% [19.3%, 20.8%]	18.3% [17.6%, 19.2%]
Current smoker, age 18-24	8,173	22.9%	18.3% [17.2%, 19.4%]	17.1% [16.4%, 17.8%]	16.7% [14.2%, 19.5%]	23.4% [19.2%, 28.3%]	NA ^b	NA
Current smoker, age 25-44	9,872	40.3%	23.5% [22.3%, 24.8%]	17.9% [17.5%, 18.4%]	20.0% [19.0%, 21.0%]	23.1% [20.7%, 25.7%]	NA	NA
Current smoker, age 45-64	7,525	42.4%	20.1% [19.0%, 21.2%]	17.8% [17.4%, 18.2%]	18.0% [17.0%, 19.0%]	21.6% [17.6%, 26.2%]	NA	NA
Current smoker, age 65+	2,763	23.9%	8.2% [7.2%, 9.3%]	7.8% [7.5%, 8.2%]	8.5% [7.7%, 9.4%]	8.4% [6.9%, 10.1%]	NA	NA
Current smoker, Hispanic	5,033	25.0%	14.4% [13.3%, 15.5%]	10.9% [10.4%, 11.5%]	11.2% [10.2%, 12.2%]	13.7% [11.4%, 16.3%]	18.5% [17.0%, 20.1%]	15.2% [13.8%, 16.6%]
Current smoker, White non-Hispanic	16,707	37.1%	19.4% [18.2%, 20.6%]	17.5% [17.2%, 17.8%]	18.2% [17.3%, 19.1%]	20.9% [17.7%, 24.5%]	23.9% [23.2%, 24.7%]	22.7% [21.9%, 23.4%]
Current smoker, other non-Hispanic	6,143	33.7%	21.8% [20.6%, 23.1%]	NA	16.2% [15.2%, 17.2%]	21.7% [18.9%, 24.8%]	21.8% [20.4%, 23.3%]	19.4% [18.0%, 20.9%]
Current every-day smoker	28,337	26.3%	14.5% [13.7%, 15.4%]	12.7% [12.4%, 12.9%]	12.8% [12.3%, 13.4%]	16.1% [14.2%, 18.2%]	NA	NA
Current some-days smoker	28,337	7.9%	4.4% [4.2%, 4.7%]	3.4% [3.3%, 3.5%]	3.9% [3.6%, 4.2%]	3.9% [3.2%, 4.7%]	NA	NA

^a NSDUH's definition of a current cigarette smoker is someone who has smoked part or all of a cigarette in the past 30 days, which is more expansive than the definition used in the other surveys. However, NSDUH contains questions on lifetime smoking and current smoking. The modified definition uses these questions to construct a measure of "current smoking" that is comparable to that of the other surveys (Ryan et al., 2012). The construction of this variable is described in Appendix A.

^b Detailed age information was not available in the public use file for NSDUH 2014.

Table 2-18. Cigarette smoking* based on Wave 2 Youth Interview respondents: IPS weights

Characteristic at Wave 2	Unweighted count	PATH Study: Unweighted percentage	PATH Study: Weighted percentage, using youth IPS weights [95% confidence interval]	Percentage from 2013-2014 NHANES [95% confidence interval]	Percentage from 2014 NSDUH [95% confidence interval]	Percentage from 2014 NYTS [95% confidence interval]
Ever tried cigarette smoking, even one or two puffs	12,148	11.8%	11.9% [11.1%, 12.8%]	18.7% [15.3%, 22.6%]	14.4% [13.6%, 15.3%]	21.7% [20.2%, 23.3%]
Ever tried smoking, male	6,208	11.8%	12.1% [11.0%, 13.2%]	19.2% [14.3%, 25.4%]	14.6% [13.5%, 15.7%]	22.5% [20.8%, 24.2%]
Ever tried smoking, female	5,911	11.7%	11.7% [10.8%, 12.8%]	18.0% [13.0%, 24.4%]	14.2% [13.1%, 15.4%]	21.0% [19.2%, 22.9%]
Ever tried smoking, age 12-13	4,150	3.9%	4.0% [3.4%, 4.7%]	5.1% [3.0%, 8.6%]	3.8% [3.1%, 4.6%]	10.3% [8.6%, 12.3%]
Ever tried smoking, age 14-17	7,998	15.8%	16.0% [14.9%, 17.1%]	25.2% [20.8%, 30.2%]	19.4% [18.2%, 20.6%]	27.4% [25.4%, 29.4%]
Have smoked in past 30 days	12,068	4.0%	4.0% [3.6%, 4.5%]	3.9% [2.7%, 5.5%]	5.0% [4.6%, 5.5%]	5.8% [5.2%, 6.5%]

* Defined as ever tried a cigarette, even one or two puffs. For comparison, an additional measure of current smoking commonly applied to youth (having smoked at all in the past 30 days) is also included in this table.

The PATH Study and NSDUH both use ACASI administration for the tobacco use questions so that the interviewer does not see responses to the questions. By contrast, TUS-CPS, NHIS, and NHANES have direct questioning by an interviewer: NHIS and NHANES are conducted in person, and TUS-CPS is conducted in person and by telephone. The contexts and purposes of these surveys also differ: CPS is a general survey on unemployment, NHIS and NHANES are general health surveys, and NSDUH is a cross-sectional survey on substance use (including tobacco use) and health, including mental health. Unlike the cross-sectional prevalence surveys, the PATH Study is designed for research purposes and uses a longitudinal cohort design to assess within-person change and between-person differences in tobacco use behaviors and health over time. Other differences among the questions used in the instruments of these different studies are outlined in Appendix A.

Table 2-17 indicates the IPS-weighted estimates of current smoking from Wave 2 of the PATH Study are most similar to estimates from NHIS and NHANES. All of the 95 percent confidence intervals for percent of current cigarette smokers constructed from the PATH Study overlap with the confidence intervals for NHIS, NHANES, or both. Estimates from TUS-CPS tend to be below the estimates from the PATH Study, NHIS, and NHANES; estimates from NSDUH tend to be above the estimates from the PATH Study, NHIS, and NHANES. No evidence was found to indicate nonresponse bias in the PATH Study with respect to current cigarette smoking behavior among adults, because the PATH Study's estimates are all within the range of estimates from comparable surveys.

Table 2-18 provides estimates from the PATH Study for two common measures of cigarette smoking prevalence among Wave 2 youth respondents compared with estimates from NHANES, NSDUH, and NYTS.¹¹ Different measures of smoking are used in this report for youth than for adults. The primary measure of cigarette smoking among youth in this report is whether the youth has ever tried smoking a cigarette, even one or two puffs (see Appendix A). Another measure is current smoking, defined as having smoked at all in the past 30 days. Both are shown in Table 2-18.

Differences in target populations and administration among the youth surveys might lead to differences in their estimates. In addition, the youth survey estimates have sampling error, as demonstrated by the confidence intervals about the estimates from the comparison surveys.

¹¹ TUS-CPS does not interview persons younger than 18 about tobacco use.

Questions and their orderings also differ among the surveys, as described in Appendix A, as do the modes of administration. The PATH Study, NHANES, and NSDUH use ACASI for the questions about tobacco use by youth, and these are administered individually in a household or mobile examination center setting. The NYTS is a pencil-and-paper survey that is self-administered in the classroom. Currivan et al. (2004) found that even when telephone ACASI was used, estimates of youth smoking prevalence were lower for a telephone survey of youth smoking than for a school-based survey of the same population (see also Fowler and Stringfellow, 2001, for a discussion of higher smoking rates in school-based surveys). In addition, school-based surveys often include students who are older than 17, which is the upper age limit for youth in the PATH Study.

The PATH Study's estimates of youth smoking are lower than comparable estimates from NHANES and NSDUH. Part of this difference may be sampling error and part may be attributable to differences among the survey wordings and administrations. Moreover, the comparison surveys are from different time periods. According to the Center for Behavioral Health Statistics and Quality (CBHSQ) (2015b), cigarette smoking among teens is dropping (from 2013 to 2014, the percentage of youth who had ever tried smoking dropped by 0.3 percentage points among 12-13 year-olds, 2.4 percentage points among 14-15 year-olds, and 2.1 percentage points among 16-17 year-olds, with similar decreases from 2012 to 2013). The lower percentages found by the PATH Study may reflect, in part, a continuation of this trend. However, some of the differences among the estimates of youth smoking prevalence may be attributable to nonresponse bias or measurement error on the part of one or more of the surveys.

2.3 Statistical Approach for Addressing Nonresponse

2.3.1 Interviews

The primary approach for addressing nonresponse is to use weight adjustments that account for differential response propensities across subgroups. This section describes the weighting adjustments used to address the areas of potential nonresponse bias identified in Section 2.2 and evaluates their effectiveness.

Among numerous sources, the handbook on household surveys by the United Nations (2005, Chapter 6) and Särndal and Lundström (2005) discuss the methods and theory of using weight adjustments for nonresponse. For Wave 1, these adjustments were done at the household level and

at the person level. The Wave 1 household-level weighting adjustments calibrated the estimates to household-level population estimates for census region and household composition and size developed from the 1-year 2013 ACS (which provides high precision and accuracy due to a large sample size and high response rate). Such weighting adjustments also correct for disparities among other characteristics that might be associated with the characteristics involved in the weighting adjustments. After accounting for household-level nonresponse, households with at least one person sampled for the PATH Study were identified, and each sampled person within a household was assigned the corresponding household weight. These weights were then adjusted to account for nonresponse to the Wave 1 interview (for adults and youth) or to non-participation in the study (for shadow youth). After this adjustment for nonresponse, the weights were calibrated using a raking process to person-level population estimates also developed from the 1-year 2013 ACS. Outlier values of the sample weights were trimmed if called for, and the weights were re-raked after any such trimming.

For Wave 2, the same general approach was used for adults, youth, and shadow youth (all as defined at Wave 1), separately. The final person-level weights assigned to Wave 1 respondents served as the initial (“base”) weights for use in developing the Wave 2 weights. These weights were adjusted to account for nonresponse to the Wave 2 interview and the resulting weights raked to control totals. However, some of these control totals were sampled-based rather than population-based and they reflected characteristics of the Wave 1 population (the baseline period for this sample).

Raking to sample-based control totals, often employed with longitudinal studies (see, for example, Brick, Le, and West (2003)) can help limit drifting from some important baseline characteristics that might arise through the applications of nonresponse adjustments over time. Lundström and Särndal (1999) provide a theoretical discussion of the use of calibrating weights to sample-based controls as well as providing empirical evidence that such calibration can serve to reduce both variance and nonresponse bias.

In terms of more details on the Wave 2 weighting, the Wave 1 final weights were adjusted for nonresponse at the person level in two stages. The first adjustment accounted for nonrespondents whose eligibility status was unknown, largely those people who were unlocatable. A second adjustment for nonresponse was undertaken to account for nonrespondents known to be eligible for the Wave 2 interview. This group of nonrespondents consisted mainly of those who were contacted but chose not to complete the Wave 2 interview.

Nonresponse adjustment cells were created separately for each of the two stages. Variables from Wave 1 data collection were candidates for contributing to the formation of nonresponse adjustment cells as were some variables based on paradata. The set of candidate variables varied somewhat by age group (adult, youth, shadow youth). Examples of the types of candidate variables are screener characteristics (e.g., household size), segment level characteristics (e.g., race percentages in segment), Wave 1 tobacco use status variables (for youth and adults), and health-based characteristics (e.g., BMI). The paradata variables included factors such as the number of contacts required to complete the Wave 1 interview. A tree-based classification method was employed to identify subgroups exhibiting differential nonresponse to the Wave 2 interview. [This general approach is described by Roth et al. (2006) and Schouten and deNoij (2005).] A description of the computation of each of the two nonresponse adjustment factors follows.

Suppose Wave 1 respondent i was assigned weight $W1_i$ for Wave 1 and was assigned to cell c_U when adjusting for nonresponse among those whose eligibility status was not ascertained. The Wave 2 weight accounting for such nonresponse is represented as $W2NRUNK_i$ and was computed as the product of the Wave 1 weight and the adjustment factor associated with cell c_U :

$$W2NRUNK_i = W1_i \times \frac{\text{sum of } W1_i \text{ weights for all Wave 1 respondents assigned to cell } c_U}{\text{sum of } W1_i \text{ weights for all Wave 1 respondents with known eligibility status at Wave 2 assigned to cell } c_U}$$

After this adjustment, all those Wave 1 respondents known to be ineligible for the Wave 2 interview were removed from the weighting process, leaving for further adjustment purposes only those known to be eligible for Wave 2, whether respondent or nonrespondent. For this set of people, the weight $W2NRUNK_i$ was further adjusted for nonresponse among those known to be eligible for the Wave 2 interview.

Let c_K represent the cell to which respondent i was assigned for the nonresponse adjustment among those known to be eligible for Wave 2. Then the final Wave 2 nonresponse-adjusted weight for respondent i , represented as $W2NRWT_i$, was computed as the product of $W2NRUNK_i$ and the adjustment factor associated with cell c_K as follows

$$W2NRWT_i = W2NRUNK_i \times \frac{\text{sum of } W2NRUNK_i \text{ weights for all those eligible for Wave 2 assigned to cell } c_K}{\text{sum of } W2NRUNK_i \text{ weights for all Wave 2 eligible respondents assigned to cell } c_K}$$

Following the adjustments for nonresponse, the weights were raked to Wave 1 control totals. Extreme weights were identified and trimmed and the weights were raked again. As mentioned earlier, some of the control totals for raking were the population-based control totals used in the Wave 1 raking while others were sampled based. An important aspect to note here is that the sample-based control totals are subject to non-negligible sampling error which should be reflected in the establishment of the replicate weights, and this was done as part of the Wave 2 weighting (specifically, as part of the computation of the replicate weights).

For continuing adult respondents, the population-based controls were based on combinations of Wave 1 census region, age, race/ethnicity, sex, and educational attainment, all used during the Wave 1 weight raking process. The sampled-based controls reflected any tobacco use and e-cigarette use reported at Wave 1 cross-classified with some of the same characteristics (e.g., sex and age). The tobacco use variable had three levels: current established user, ever user but not current established user, and never user; those with missing data for tobacco use were pooled for raking purposes with those associated with the category “ever user but not current established user.”¹² The e-cigarette variable had two levels: never user and not never user; those with missing data for e-cigarette use were pooled with those associated with the category “not never user.”

For continuing youth respondents and aged-up adult respondents, the raking was done using Wave 1 census region, single-year of age, race/ethnicity, and sex (all used during the Wave 1 weight raking process), as well as any tobacco use (for respondents who were age 16 or 17 at Wave 1) and e-cigarette use (for respondents who were age 17 at Wave 1);¹³ the tobacco use and e-cigarette use variables were both defined using two levels: never user and not never user; those with missing data were pooled with those associated with the category “not never user.”

For aged-up youth respondents, the raking was done using Wave 1 census region, single-year of age, race/ethnicity, and sex.

¹² For the purpose of sample-based raking, cases with missing values were pooled with other cases in such a way that the categories thought to be of most analytic interest were unaffected.

¹³ The ages for which the e-cigarette use and any tobacco use dimensions were used for raking were determined by the need to have sufficient sample sizes for stability of the weighting adjustments.

A trimming procedure was implemented as called for, after the raking was completed. Whenever trimming was implemented, the corresponding set of weights was re-raked. After the raking and trimming process, the final weight, denoted as $W2RKWT_i$, was computed for each participant as

$$W2RKWT_i = W2NRWT_i \times RT_i ,$$

where RT_i is the combined raking and trimming adjustment factor associated with respondent i .

Replicate weights were created for variance estimation purposes using balanced repeated replication methodology, reflecting the sample design and weighting process including raking to control totals that included some PATH Study sample estimates, as described above.

Tables 2-19 through 2-25 repeat the analyses in Tables 2-8 through 2-14; however, the Wave 2 nonrespondent estimates are replaced by estimates for the Wave 1 respondents who remained eligible at Wave 2, and the Wave 2 respondent estimates use the Wave 2 final weights (rather than the Wave 1 final weights). Differences between the weighted estimates for Wave 2 and Wave 1 respondents therefore reflect the extent to which the Wave 2 weighting process corrected for potential nonresponse bias. Across all seven tables, the biggest difference found was 0.8 percent (for non-Hispanic White aged-up adults, see Table 2-24) and this was not statistically significant at the .05 level. These results suggest that the Wave 2 weighting nonresponse adjustments were highly effective.

Table 2-19 shows that the Wave 2 weighting process corrected for potential nonresponse bias associated with Wave 1 demographic characteristics and health insurance coverage among continuing adults at Wave 2. All the differences between the point estimates are essentially zero. The astute reader will notice that the 95 percent confidence intervals around the difference estimates barely exclude zero for a small number of subgroups (males and 18-44 year-olds), however this is primarily due to the large sample sizes.¹⁴

¹⁴ Part of the explanation also lies in technical subtleties associated with the Wave 2 weighting process. Persons ineligible at Wave 2 were included with the Wave 2 respondents when raking to Wave 1 demographics and tobacco use estimates but are excluded from the estimates in the table. Also, cases with missing values for the tobacco use estimates used in the Wave 2 raking process were pooled with cases with non-missing values (as described above) during raking whereas they are excluded from the estimates in the table.

Table 2-19. Comparison of Wave 2 continuing Adult Interview respondents with Wave 1 Adult Interview respondents

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 1 adult respondents ^c		Difference in weighted percentages [Wave 2 – Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Sex					
Male	13,080	47.9% [47.3%, 48.5%]	16,147	47.9% [47.4%, 48.5%]	-0.0% [-0.0%, -0.0%]
Female	13,356	52.1% [51.5%, 52.7%]	15,881	52.1% [51.5%, 52.6%]	0.0% [0.0%, 0.0%]
Age group					
18-24	7,330	13.0% [12.6%, 13.5%]	9,072	13.1% [12.7%, 13.4%]	-0.0% [-0.0%, -0.0%]
25-44	9,233	34.4% [33.8%, 35.0%]	11,216	34.4% [33.9%, 34.9%]	-0.0% [-0.0%, -0.0%]
45-64	7,375	34.6% [34.0%, 35.2%]	8,731	34.6% [34.0%, 35.1%]	0.0% [-0.0%, 0.0%]
65+	2,516	18.0% [17.5%, 18.5%]	3,027	18.0% [17.6%, 18.4%]	0.0% [-0.0%, 0.1%]
Race/ethnicity					
Non-Hispanic White alone	15,767	65.9% [65.4%, 66.5%]	19,138	66.0% [65.5%, 66.5%]	-0.0% [-0.1%, 0.0%]
Other	10,252	34.1% [33.5%, 34.6%]	12,368	34.0% [33.5%, 34.5%]	0.0% [-0.0%, 0.1%]
Health insurance					
Yes	21,269	85.4% [84.8%, 86.1%]	25,550	85.5% [84.9%, 86.0%]	-0.0% [-0.2%, 0.2%]
No	4,924	14.6% [13.9%, 15.2%]	6,110	14.5% [14.0%, 15.1%]	0.0% [-0.2%, 0.2%]
Education					
< HS or GED	5,279	16.6% [16.2%, 17.1%]	6,376	16.6% [16.2%, 17.0%]	-0.0% [-0.0%, 0.0%]
HS	6,056	24.3% [23.8%, 24.8%]	7,477	24.3% [23.8%, 24.7%]	0.0% [-0.0%, 0.1%]
Some college, no degree	9,308	31.1% [30.6%, 31.7%]	11,231	31.1% [30.6%, 31.6%]	0.0% [-0.0%, 0.1%]
Bachelor degree +	5,684	28.0% [27.4%, 28.5%]	6,771	28.0% [27.5%, 28.5%]	-0.0% [-0.1%, 0.0%]

Table 2-19. Comparison of Wave 2 continuing Adult Interview respondents with Wave 1 Adult Interview respondents (continued)

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 1 adult respondents ^c		Difference in weighted percentages [Wave 2 - Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Tobacco use status ^b					
Current established user	11,693	23.8% [23.2%, 24.4%]	14,234	23.9% [23.3%, 24.5%]	-0.1% [-0.1%, -0.0%]
Not current established user	14,039	76.2% [75.6%, 76.8%]	16,903	76.1% [75.5%, 76.7%]	0.1% [0.0%, 0.1%]

- ^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.
- ^b A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.
- ^c Excludes those who were ineligible at Wave 2.

Table 2-20. Comparison of Wave 1 tobacco use* rates for Wave 2 continuing Adult Interview respondents with Wave 1 Adult Interview respondents

Characteristic at Wave 1 ^a	Wave 2 continuing adult respondents		Wave 1 adult respondents ^b		Difference in weighted percentages [Wave 2 – Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Overall	25,732	23.8% [23.2%, 24.4%]	31,137	23.9% [23.3%, 24.5%]	-0.1% [-0.1%, -0.0%]
Sex					
Male	12,738	29.6% [28.8%, 30.5%]	15,699	29.8% [28.9%, 30.6%]	-0.2% [-0.3%, -0.1%]
Female	12,973	18.4% [17.7%, 19.1%]	15,412	18.4% [17.7%, 19.1%]	-0.0% [-0.1%, 0.0%]
Age group					
18-24	7,214	29.0% [27.6%, 30.4%]	8,922	29.0% [27.6%, 30.4%]	-0.0% [-0.1%, 0.0%]
25-44	9,052	28.8% [27.8%, 29.8%]	10,990	29.0% [28.1%, 29.9%]	-0.2% [-0.5%, 0.1%]
45-64	7,113	23.2% [22.2%, 24.2%]	8,406	23.2% [22.3%, 24.1%]	-0.1% [-0.4%, 0.2%]
65+	2,348	10.8% [9.6%, 12.1%]	2,808	10.7% [9.6%, 11.9%]	0.1% [-0.3%, 0.5%]
Race/ethnicity					
Non-Hispanic White alone	15,428	25.2% [24.3%, 26.1%]	18,700	25.3% [24.4%, 26.2%]	-0.1% [-0.2%, -0.0%]
Other	9,923	21.0% [20.2%, 21.8%]	11,962	21.1% [20.4%, 21.9%]	-0.1% [-0.2%, 0.0%]

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Excludes those who were ineligible at Wave 2.

* Defined as current established use of tobacco. A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 2-21. Comparison of Wave 2 continuing Youth Interview respondents with Wave 1 youth respondents who were continuing youth at Wave 2

Characteristic at Wave 1 ^a	Wave 2 continuing youth respondents		Wave 1 youth respondents ^c		Difference in weighted percentages [Wave 2 - Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	5,164	51.4% [50.4%, 52.4%]	5,829	51.3% [50.4%, 52.3%]	0.0% [-0.0%, 0.1%]
Female	4,892	48.6% [47.6%, 49.6%]	5,533	48.7% [47.7%, 49.6%]	-0.0% [-0.1%, 0.0%]
Age group					
12-13	4,164	40.7% [39.8%, 41.7%]	4,678	40.9% [40.0%, 41.8%]	-0.2% [-0.2%, -0.1%]
14-17	5,917	59.3% [58.3%, 60.2%]	6,718	59.1% [58.2%, 60.0%]	0.2% [0.1%, 0.2%]
Race/ethnicity					
Non-Hispanic White alone	4,731	54.5% [53.5%, 55.5%]	5,347	54.5% [53.5%, 55.4%]	0.0% [-0.1%, 0.1%]
Other	5,141	45.5% [44.5%, 46.5%]	5,808	45.5% [44.6%, 46.5%]	-0.0% [-0.1%, 0.1%]
Tobacco use status^b					
Ever user	1,665	17.1% [16.3%, 17.9%]	1,926	17.3% [16.4%, 18.3%]	-0.2% [-0.7%, 0.3%]
Never user	8,006	82.9% [82.1%, 83.7%]	8,988	82.7% [81.7%, 83.6%]	0.2% [-0.3%, 0.7%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

^c Excludes those who were ineligible at Wave 2 and also excludes Wave 1 youth who were aged-up adults at Wave 2.

Table 2-22. Comparison of Wave 1 tobacco use* rates for Wave 2 continuing Youth Interview respondents with Wave 1 youth respondents who were continuing youth at Wave 2

Characteristic at Wave 1 ^a	Wave 2 continuing youth respondents		Wave 1 youth respondents ^b		Difference in weighted percentages [Wave 2 - Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Overall	9,671	17.1% [16.3%, 17.9%]	10,914	17.3% [16.4%, 18.3%]	-0.2% [-0.7%, 0.3%]
Sex					
Male	4,922	18.1% [17.1%, 19.2%]	5,541	18.4% [17.3%, 19.6%]	-0.3% [-1.0%, 0.3%]
Female	4,727	16.1% [15.1%, 17.2%]	5,343	16.2% [15.2%, 17.4%]	-0.1% [-0.7%, 0.4%]
Age group					
12-13	3,923	7.8% [6.8%, 8.9%]	4,396	7.8% [6.9%, 8.9%]	-0.1% [-0.4%, 0.2%]
14-17	5,748	23.3% [22.3%, 24.5%]	6,517	23.7% [22.5%, 24.9%]	-0.3% [-1.1%, 0.5%]
Race/ethnicity					
Non-Hispanic White alone	4,555	18.0% [16.9%, 19.1%]	5,138	18.4% [17.1%, 19.8%]	-0.4% [-1.0%, 0.2%]
Other	4,930	16.3% [15.3%, 17.4%]	5,563	16.3% [15.2%, 17.4%]	0.0% [-0.6%, 0.6%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Excludes those who were ineligible at Wave 2 and also excludes Wave 1 youth who were aged-up adults at Wave 2.

* Defined as ever use of tobacco. An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-23. Comparison of Wave 2 aged-up Adult Interview respondents with Wave 1 youth respondents who were aged-up adults at Wave 2

Characteristic at Wave 1 ^a	Wave 2 aged-up adult respondents		Wave 1 youth respondents ^c		Difference in weighted percentages [Wave 2 – Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	957	50.7% [48.5%, 53.0%]	1,126	51.0% [48.9%, 53.0%]	-0.2% [-0.6%, 0.2%]
Female	956	49.3% [47.0%, 51.5%]	1,102	49.0% [47.0%, 51.1%]	0.2% [-0.2%, 0.6%]
Race/ethnicity					
Non-Hispanic White alone	954	55.1% [52.8%, 57.3%]	1,127	55.3% [53.2%, 57.4%]	-0.2% [-0.6%, 0.2%]
Other	945	44.9% [42.7%, 47.2%]	1,086	44.7% [42.6%, 46.8%]	0.2% [-0.2%, 0.6%]
Tobacco use status^b					
Ever user	811	43.2% [40.9%, 45.4%]	937	42.6% [40.3%, 44.9%]	0.6% [-1.7%, 2.8%]
Never user	1,061	56.8% [54.6%, 59.1%]	1,241	57.4% [55.1%, 59.7%]	-0.6% [-2.8%, 1.7%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

^c Excludes those who were ineligible at Wave 2 and also excludes Wave 1 youth who were continuing youth at Wave 2.

Table 2-24. Comparison of Wave 1 tobacco use* rates for Wave 2 aged-up Adult Interview respondents with Wave 1 youth respondents who were aged-up adults at Wave 2

Characteristic at Wave 1 ^a	Wave 2 aged-up adult respondents		Wave 1 youth respondents ^b		Difference in weighted percentages [Wave 2 – Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Overall	1,872	43.2% [40.9%, 45.4%]	2,178	42.6% [40.3%, 44.9%]	0.6% [-1.7%, 2.8%]
Sex					
Male	933	46.6% [43.4%, 49.8%]	1,097	46.1% [42.9%, 49.4%]	0.5% [-2.6%, 3.5%]
Female	936	39.7% [36.6%, 42.9%]	1,078	39.0% [36.0%, 42.0%]	0.7% [-2.2%, 3.6%]
Race/ethnicity					
Non-Hispanic White alone	936	47.0% [43.8%, 50.2%]	1,105	46.2% [42.8%, 49.7%]	0.8% [-2.3%, 3.8%]
Other	919	38.2% [35.1%, 41.4%]	1,055	37.9% [34.4%, 41.4%]	0.3% [-2.8%, 3.5%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Excludes those who were ineligible at Wave 2 and also excludes Wave 1 youth who were continuing youth at Wave 2.

* Defined as ever use of tobacco. An ‘ever user’ is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 2-25. Comparison of Wave 2 aged-up Youth Interview respondents with Wave 1 shadow youth who were aged-up youth at Wave 2

Characteristic at Wave 1 ^a	Wave 2 aged-up youth respondents		Wave 1 shadow youth ^b		Difference in weighted percentages [Wave 2 - Wave 1] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 2 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using shadow youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	1,055	50.3% [48.1%, 52.4%]	1,292	50.2% [48.3%, 52.1%]	0.1% [-0.3%, 0.4%]
Female	1,036	49.7% [47.6%, 51.9%]	1,258	49.8% [47.9%, 51.7%]	-0.1% [-0.4%, 0.3%]
Race/ethnicity					
Non-Hispanic White alone	1,008	52.1% [50.0%, 54.3%]	1,223	52.2% [50.2%, 54.1%]	-0.0% [-0.4%, 0.3%]
Other	1,078	47.9% [45.7%, 50.0%]	1,321	47.8% [45.9%, 49.8%]	0.0% [-0.3%, 0.4%]

^a The characteristics are as reported in the Household Screener at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Excludes those who were ineligible at Wave 2 and also excludes Wave 1 shadow youth who were continuing shadow youth at Wave 2.

Table 2-20 demonstrates the effectiveness of the Wave 2 weighting process on Wave 1 tobacco use estimates for continuing adults. The differences between the point estimates are not substantive, despite the 95 percent confidence intervals indicating that estimates overall and for males and non-Hispanic Whites are marginally lower among Wave 2 respondents. The confidence intervals around the estimates of difference between Wave 1 and Wave 2 respondents are narrow reflecting high correlation between the two groups over time and the use of sample-based raking to Wave 1 tobacco use estimates.

Among continuing youth, using the Wave 2 final weights, Table 2-21 shows that 12-13 year-olds are slightly underrepresented among Wave 2 respondents compared to Wave 1 respondents. However, the magnitude of the difference is not practically meaningful. Table 2-22 shows no evidence of nonresponse bias for Wave 1 estimates of ever tobacco use among continuing youth at Wave 2.

Similarly, no evidence of potential nonresponse bias was found for aged-up adults or aged-up youth based on the estimates in Tables 2-23, 2-24, and 2-25.

Estimates of Wave 2 adult cigarette smoking prevalence in Table 2-26 using the Wave 2 final weights are similar to the estimates using the Wave 1 IPS weights; both are in the range of values obtained by other surveys. The use of the Wave 2 weights resulted in a slight decrease in estimated cigarette smoking prevalence for females and non-White non-Hispanics.

Table 2-27 examines the effect of the Wave 2 final weights on estimates calculated for youth. Cigarette smoking prevalence estimates with the Wave 1 IPS weights and with the Wave 2 weights were generally lower than estimates from other surveys although, as noted above, the surveys took place in different time periods.

Table 2-26. Current cigarette smoking based on Wave 2 Adult Interview respondents

Characteristic at Wave 2	Un-weighted count	PATH Study: Unweighted percentage	PATH Study: Weighted percentage, using adult IPS weights [95% confidence interval]	PATH Study: Weighted percentage, using adult Wave 2 final weights [95% confidence interval]	Percentage from 2010-2011 TUS-CPS [95% confidence interval]	Percentage from 2014 NHIS [95% confidence interval]	Percentage from 2013-2014 NHANES [95% confidence interval]	Percentage from 2014 NSDUH, original definition ^a [95% confidence interval]	Percentage from 2014 NSDUH, modified definition ^a [95% confidence interval]
Current smoker	28,337	34.2%	19.0% [18.1%, 19.9%]	18.6% [18.1%, 19.2%]	16.1% [15.8%, 16.3%]	16.7% [16.1%, 17.4%]	20.0% [17.8%, 22.3%]	22.7% [22.1%, 23.2%]	20.9% [20.3%, 21.4%]
Current smoker, male	14,014	35.3%	21.2% [20.2%, 22.3%]	21.3% [20.6%, 22.0%]	18.0% [17.7%, 18.4%]	18.8% [18.0%, 19.7%]	21.6% [19.4%, 24.1%]	25.5% [24.7%, 26.3%]	23.6% [22.8%, 24.4%]
Current smoker, female	14,297	33.3%	17.2% [16.2%, 18.1%]	16.2% [15.6%, 16.8%]	14.2% [13.9%, 14.5%]	14.8% [13.9%, 15.7%]	18.4% [15.6%, 21.6%]	20.1% [19.3%, 20.8%]	18.3% [17.6%, 19.2%]
Current smoker, age 18-24	8,173	22.9%	18.3% [17.2%, 19.4%]	18.4% [17.4%, 19.4%]	17.1% [16.4%, 17.8%]	16.7% [14.2%, 19.5%]	23.4% [19.2%, 28.3%]	NA ^b	NA
Current smoker, age 25-44	9,872	40.3%	23.5% [22.3%, 24.8%]	23.6% [22.6%, 24.6%]	17.9% [17.5%, 18.4%]	20.0% [19.0%, 21.0%]	23.1% [20.7%, 25.7%]	NA	NA
Current smoker, age 45-64	7,525	42.4%	20.1% [19.0%, 21.2%]	19.7% [18.8%, 20.6%]	17.8% [17.4%, 18.2%]	18.0% [17.0%, 19.0%]	21.6% [17.6%, 26.2%]	NA	NA
Current smoker, age 65+	2,763	23.9%	8.2% [7.2%, 9.3%]	8.0% [7.0%, 9.1%]	7.8% [7.5%, 8.2%]	8.5% [7.7%, 9.4%]	8.4% [6.9%, 10.1%]	NA	NA
Current smoker, Hispanic	5,033	25.0%	14.4% [13.3%, 15.5%]	14.5% [13.6%, 15.5%]	10.9% [10.4%, 11.5%]	11.2% [10.2%, 12.2%]	13.7% [11.4%, 16.3%]	18.5% [17.0%, 20.1%]	15.2% [13.8%, 16.6%]
Current smoker, White non-Hispanic	16,707	37.1%	19.4% [18.2%, 20.6%]	19.4% [18.6%, 20.1%]	17.5% [17.2%, 17.8%]	18.2% [17.3%, 19.1%]	20.9% [17.7%, 24.5%]	23.9% [23.2%, 24.7%]	22.7% [21.9%, 23.4%]
Current smoker, other non-Hispanic	6,143	33.7%	21.8% [20.6%, 23.1%]	19.3% [18.3%, 20.3%]	NA	16.2% [15.2%, 17.2%]	21.7% [18.9%, 24.8%]	21.8% [20.4%, 23.3%]	19.4% [18.0%, 20.9%]
Current every-day smoker	28,337	26.3%	14.5% [13.7%, 15.4%]	14.2% [13.7%, 14.6%]	12.7% [12.4%, 12.9%]	12.8% [12.3%, 13.4%]	16.1% [14.2%, 18.2%]	NA	NA
Current some-days smoker	28,337	7.9%	4.4% [4.2%, 4.7%]	4.5% [4.2%, 4.7%]	3.4% [3.3%, 3.5%]	3.9% [3.6%, 4.2%]	3.9% [3.2%, 4.7%]	NA	NA

^a NSDUH's definition of a current cigarette smoker is someone who has smoked part or all of a cigarette in the past 30 days, which is more expansive than the definition used in the other surveys. However, NSDUH contains questions on lifetime smoking and current smoking. The modified definition uses these questions to construct a measure of "current smoking" that is comparable to that of the other surveys (Ryan et al., 2012). The construction of this variable is described in Appendix A.

^b Detailed age information was not available in the public use file for NSDUH 2014.

Table 2-27. Cigarette smoking* based on Wave 2 Youth Interview respondents

Characteristic at Wave 2	Unweighted count	PATH Study: Unweighted percentage	PATH Study: Weighted percentage, using youth IPS weights [95% confidence interval]	PATH Study: Weighted percentage, using youth Wave 2 final weights [95% confidence interval]	Percentage from 2013-2014 NHANES [95% confidence interval]	Percentage from 2014 NSDUH [95% confidence interval]	Percentage from 2014 NYTS [95% confidence interval]
Ever tried cigarette smoking, even one or two puffs	12,148	11.8%	11.9% [11.1%, 12.8%]	11.7% [11.0%, 12.3%]	18.7% [15.3%, 22.6%]	14.4% [13.6%, 15.3%]	21.7% [20.2%, 23.3%]
Ever tried smoking, male	6,208	11.8%	12.1% [11.0%, 13.2%]	11.8% [10.9%, 12.7%]	19.2% [14.3%, 25.4%]	14.6% [13.5%, 15.7%]	22.5% [20.8%, 24.2%]
Ever tried smoking, female	5,911	11.7%	11.7% [10.8%, 12.8%]	11.5% [10.7%, 12.4%]	18.0% [13.0%, 24.4%]	14.2% [13.1%, 15.4%]	21.0% [19.2%, 22.9%]
Ever tried smoking, age 12-13	4,150	3.9%	4.0% [3.4%, 4.7%]	3.8% [3.2%, 4.4%]	5.1% [3.0%, 8.6%]	3.8% [3.1%, 4.6%]	10.3% [8.6%, 12.3%]
Ever tried smoking, age 14-17	7,998	15.8%	16.0% [14.9%, 17.1%]	15.7% [14.9%, 16.5%]	25.2% [20.8%, 30.2%]	19.4% [18.2%, 20.6%]	27.4% [25.4%, 29.4%]
Have smoked in past 30 days	12,068	4.0%	4.0% [3.6%, 4.5%]	4.0% [3.6%, 4.4%]	3.9% [2.7%, 5.5%]	5.0% [4.6%, 5.5%]	5.8% [5.2%, 6.5%]

* Defined as ever tried a cigarette, even one or two puffs. For comparison, an additional measure of current smoking commonly applied to youth (having smoked at all in the past 30 days) is also included in this table.

2.3.2 Biospecimens

The results in Tables 2-15 and 2-16 show estimates from the aged-up adults who provided urine or blood specimens, respectively, at Wave 2, but not all specimens collected will be analyzed in the laboratory. The biospecimens chosen initially for laboratory analysis are likely to come from a probability sample of aged-up adults who are in specified tobacco use categories.¹⁵ In other categories of tobacco use, no samples may be selected for laboratory analysis. Consequently, the samples of biospecimens from Wave 2 aged-up adults that are analyzed may not be representative of the population of adults age 18 as a whole. If desired, an additional set of nonresponse-adjusted weights can be developed for these adults, following procedures similar to those described in the memo on Wave 1 biospecimen weighting procedures submitted to OMB (and approved on October 9, 2015). These weighting adjustments would address issues such as the overrepresentation of current established tobacco users among aged-up adults who provided urine and/or blood specimens, and the selection of adults whose biospecimens are sent for laboratory analysis.

2.4 Summary of Findings

Response Rates

As reported in Section 2.1, the response rates¹⁶ for the PATH Study Wave 2 interviews were lower than projected for continuing adults and for continuing and aged-up youth, but higher than projected for aged-up adults (see Table 2-28). The unweighted response rates for the biospecimen collections in Wave 2 were all higher than projected.

Table 2-28. Summary of PATH Study Wave 2 response rates

Group	Unweighted response rate	Weighted response rate	Projected response rate*
Continuing adults, Adult Interview	82.6%	83.1%	86%
Continuing youth, Youth Interview	88.5%	88.4%	90%
Aged-up adults, Adult Interview	85.9%	85.7%	85%
Aged-up youth, Youth Interview	82.0%	82.1%	88%
Continuing adults, urine collection	96.4%	-	80%
Aged-up adults, urine collection	82.8%	-	69%
Aged-up adults, blood collection	47.4%	-	45%

* Provided in the Revision Request to OMB for Wave 2 data and biospecimen collections.

¹⁵ Urine specimens from some continuing adults at Wave 2 will also be chosen for laboratory analysis.

¹⁶ The weighted response rates were computed using inverse probability of selection weights.

The differential weighted response rates to the Wave 2 interview were modest for tobacco use status and demographic subgroups (see Tables 2-2 to 2-5). Unweighted response rates to the urine collection were consistently high among subgroups of continuing adults (see Table 2-6). The largest differential unweighted response rates were for the current established tobacco use status of aged-up adults asked to provide urine and blood specimens: response rates for current established users were about eight percentage points higher for both types of biospecimen than for other aged-up adults (see Table 2-7), which suggests a heightened potential for nonresponse bias.

Nonresponse Bias Analysis

Nonresponse bias analysis indicates that estimates of many key demographic and Wave 1 tobacco use variables calculated using the Wave 1 final weights are comparable for Wave 2 respondents and nonrespondents. However, males, 18-24 year-olds, and those with high school education are underrepresented among continuing adult respondents; and 45-64 year-olds, persons with health insurance, and those with at least a bachelor's degree are overrepresented among respondents (see Table 2-8). Estimates of current established tobacco use are lower overall, and for males, 18-44 year-olds, and non-Hispanic Whites among continuing adult respondents (see Table 2-9). Estimates of ever use of tobacco are lower overall, and for females, 14-17 year-olds, and non-Hispanic Whites among continuing youth respondents compared to nonrespondents (see Table 2-11). No evidence of potential nonresponse bias was found for aged-up adults or aged-up youth (see Tables 2-12 to 2-14). Based on these results, for some subgroups, the PATH Study may experience attrition patterns that are similar to those in other longitudinal surveys. Cunradi et al. (2005) and Young et al. (2006) have found that smokers were less likely to be retained in subsequent waves of surveys than nonsmokers.

Results of the Wave 2 biospecimen nonresponse bias analysis for aged-up adults found no evidence of nonresponse bias with respect to sex and race/ethnicity; however, current established tobacco users were overrepresented among both urine and blood specimen providers. No nonresponse bias analysis was necessary for urine collection from continuing adults due to the high response rate of 96 percent (see Table 2-6).

When compared to national cross-sectional surveys that measure tobacco use (TUS-CPS, NHIS, NHANES, and NSDUH), estimates of adult cigarette smoking from the PATH Study Wave 2 sample are roughly in the middle of the range of estimates on smoking. There is no indication of nonresponse bias with respect to this measure.

PATH Study estimates of the selected youth cigarette smoking measure from the full Wave 2 sample are at the low end of estimates in comparison with national cross-sectional surveys that measure tobacco use (NHANES, NSDUH, and NYTS). However, estimates from the comparison surveys are from 2013 through 2014 while those from the PATH Study are from October 2014 through October 2015, and evidence suggests the use of traditional cigarettes is declining among youth. The difference among surveys on time period alone is not large enough to account for the different estimates; as indicated in Section 2.2.2, time period is one of a number of factors that may explain the different estimates.

Statistical Approach for Addressing Nonresponse

The approach used to reduce potential nonresponse bias in Wave 2 of the PATH Study was to adjust the Wave 1 final weights of respondents at the adult and youth levels to account for nonrespondents. Results of applying this approach to the full Wave 2 sample indicate the nonresponse adjustments essentially eliminated discrepancies between estimates based on Wave 1 respondents and estimates based on Wave 2 respondents with respect to demographic characteristics and Wave 1 tobacco use, for all age groups.

Procedures similar to those described in the memo on Wave 1 biospecimen weighting procedures submitted to OMB (and approved on October 9, 2015) can be used to address nonresponse among adults asked to provide biospecimens at Wave 2.

Estimates of adult cigarette smoking at Wave 2 using the Wave 1 IPS weights (before any nonresponse adjustments) are in line with estimates from other surveys; agreement in these estimates was preserved using the Wave 2 final weights. Weighting adjustments for youth had little effect on the Wave 2 estimates of youth cigarette smoking.

Wave 3 of the PATH Study is at the approximate mid-point of data and biospecimen collections. This section discusses the predicted response rates for Wave 3, an interim nonresponse analysis, and the study's planned statistical approach for addressing nonresponse in Wave 3. All such analyses are based on data collected from the cases in replicate group 1 that had been released by April 29, 2016.

All study participants who completed an adult interview in Wave 1 or Wave 2 are continuing adults and, if eligible, are asked to complete an Adult Interview in Wave 3. Study participants whose last completed interview prior to Wave 3 was a Youth Interview are aged-up adults in Wave 3 if they completed a Wave 3 Adult Interview, and are continuing youth if they completed a Wave 3 Youth Interview. Wave 2 shadow youth who completed a Wave 3 Youth Interview are aged-up youth. Nonrespondents and interim cases for Wave 3, however, do not have a Wave 3 interview date, so the following procedure was used to determine their ages and participant types for this report.¹⁷ A Wave 2 youth who has not responded in Wave 3 is classified as a continuing youth if his/her age was determined to be 17 or younger on the age classification date; otherwise he/she is classified as an aged-up adult.¹⁸ A similar classification rule was used for persons who were shadow youth at Wave 2. A Wave 2 shadow youth who has not responded in Wave 3 is classified as an aged-up youth if he/she was determined to have attained age 12 on or before the age classification date.

As stated in Section 1, the PATH Study Wave 1 sample was divided among four replicate groups. Replicate group 1, which consisted of the addresses that were released to the field in September 2013, obtained Wave 1 responses from 5,951 adults and 2,698 youth, and parental consent for 1,414 shadow youth. Approximately 96 percent of those cases had been released to the field for Wave 3 as of April 29, 2016. Replicate group 1 roughly corresponds to the earliest set of follow-ups in Wave 3.

¹⁷ Please refer to Section 2.1 for a general explanation of how the PATH Study sets an age classification date to determine the age of a nonrespondent in a follow-up wave for reporting purposes.

¹⁸ Under this rule, for retention rate calculation purposes, nonresponding and interim youth in Wave 3 who were youth respondents in Wave 2 are assigned to the category (continuing youth or aged-up adult) that would result if they completed an interview on their Wave 3 age classification date.

Table 3-1 displays the status of cases from replicate group 1 that had been released to the field as of April 29, 2016 for the four categories of continuing adults, continuing youth, aged-up adults, and aged-up youth. The finalized cases include respondents and finalized nonrespondents of all types (eligible, ineligible, and unknown eligibility).

Table 3-1. Status of Wave 3 released cases from replicate group 1, as of April 29, 2016

Group	Case status					
	Finalized		Interim		Total	
	n	%	n	%	n	%
Continuing adults	5,154	84.7	934	15.3	6,088	100.0
Continuing youth	1,832	86.0	297	14.0	2,129	100.0
Aged-up adults	405	81.2	94	18.8	499	100.0
Aged-up youth	440	83.2	89	16.8	529	100.0

3.1 Predicted Response Rates

This section summarizes the three types of predicted response rate calculations used for Wave 3 of the PATH Study: the interim retention and recruitment rates for the interviews and the interim response rates for the biospecimen collections. In Section 3.1.1, retention rates for Wave 3 apply to persons who completed the Adult Interview in Wave 2 (i.e., continuing adults), and persons who completed the Youth Interview in Wave 2 and who are age 17 or younger at Wave 3 (i.e., continuing youth). In Section 3.1.2, recruitment rates for Wave 3 apply to those Wave 2 respondents who have aged up, either as Wave 2 shadow youth who have turned age 12 and are eligible to participate in the Wave 3 Youth Interview (i.e., aged-up youth), or as Wave 2 youth who have turned age 18 and are eligible to participate in the Wave 3 Adult Interview (i.e., aged-up adults). The Wave 3 predicted recruitment rates and the Wave 3 predicted retention rates reported in these sections are conditional on Wave 2 response. For example, the denominator for calculating the Wave 3 predicted retention rate for continuing adults is the count of Wave 2 Adult Interview respondents who are eligible for Wave 3 Adult Interview. Section 3.1.3 presents the predicted recruitment and retention rates among all Wave 1 respondents that are obtained by combining the prediction results for Wave 2 respondents with predicted numbers of completed Wave 3 interviews among the Wave 2 nonrespondents. Response rates for biospecimen collections appear in Section 3.1.4 and refer to the percentages of persons providing biospecimens among those who are asked to provide biospecimens.

3.1.1 Predicted Retention Rates among Wave 2 Respondents for Continuing Adults and Continuing Youth

This section reports retention rates for continuing adults who completed the Adult Interview at Wave 2, and continuing youth who completed the Youth Interview at Wave 2 and remained eligible for the Youth Interview at Wave 3.

Method

Consistent with the response rate calculation guidelines specified by the Office of Management and Budget (2006), final retention rates for Wave 3 will be calculated for adults as the ratio of the number of Wave 3 Adult Interview completed cases (or sufficient partials) to the number of cases eligible for the Wave 3 Adult Interview. A simplified formulation will be used for this report because the eligibility status of some interim cases is unknown. The simplified formulation corresponds to AAPOR RR1 (AAPOR, 2015), which treats all completed cases from Wave 2 as eligible for Wave 3. The predicted retention rates are therefore slightly conservative because some ineligible persons are included in the denominator.

If all the Wave 3 cases were finalized, the RR1 retention rate for continuing adults would be calculated as (number of completes or sufficient partials at Wave 3)/(number of Wave 2 completed cases minus number of persons who died, were in a correctional facility, or left the country), where the denominator can equivalently be expressed as the sum of the respondents and finalized eligible nonrespondents. Because the PATH Study Wave 3 data collection is ongoing, however, the formula must consider “nonfinalized” or interim status cases as well as finalized cases; in this sense, the retention and recruitment rates presented in this Interim Report are “predicted.” Furthermore, the cases in replicate group 1 that had not been released by April 29, 2016 are not included in these analyses – these cases represent study participants who took relatively longer to complete a Wave 2 interview (which in some instances was because they were reluctant to respond), and who therefore may be less inclined to respond at Wave 3. To help offset this selection bias among the analyzed cases, for prediction purposes, the interim refusals¹⁹ and persons who are difficult to locate were

¹⁹ Interim refusals for adults are cases that initially declined to participate in the Wave 3 interview but are still being followed for refusal conversion attempts. Interim refusals for youth are those for whom the parent initially declined permission for the youth to participate in the Wave 3 interview, but whose parents are still being followed for refusal conversion attempts.

considered to be finalized nonrespondents. In this report, the unweighted retention rate for continuing adults is calculated as

$$RRU_{CA} = (C_{CA} + \sum_{i=1}^{I_{CA}} \hat{p}_{i,CA}) / (C_{CA} + N_{CA} + I_{CA}),$$

where

- C_{CA} = number of Wave 3 completed cases or sufficient partials among Wave 2 adult respondents;
- N_{CA} = number of Wave 3 finalized nonrespondents among Wave 2 adult respondents;
- I_{CA} = number of Wave 3 interim cases among Wave 2 adult respondents; and
- $\hat{p}_{i,CA}$ = predicted probability of interim continuing adult i becoming a respondent.

For continuing youth, the denominator of the response rate is defined using the age classification date described earlier, and the unweighted retention rate is calculated as

$$RRU_{CY} = (C_{CY} + \sum_{i=1}^{I_{CY}} \hat{p}_{i,CY}) / (C_{CY} + N_{CY} + I_{CY}),$$

where

- C_{CY} = number of Wave 3 completed Youth Interviews or sufficient partials among Wave 2 youth respondents;
- N_{CY} = number of Wave 3 finalized nonrespondents among Wave 2 youth respondents who were age 17 or younger on the age classification date;
- I_{CY} = number of Wave 3 interim cases among Wave 2 youth respondents who were age 17 or younger on the age classification date; and
- $\hat{p}_{i,CY}$ = predicted probability of interim continuing youth i becoming a respondent.

The weighted response rates are computed similarly, with each count of finalized cases replaced by the sum of the Wave 1 IPS weights (AIPSWT or YIPSWT) for individuals in that category. The sums of the predicted probabilities for interim cases are replaced by $\sum_{i=1}^{I_{CA}} AIPSWT_i * \hat{p}_{i,CA}$ in the formula for RRU_{CA} and by $\sum_{i=1}^{I_{CY}} YIPSWT_i * \hat{p}_{i,CY}$, in the formula for RRU_{CY} .

The probability that an interim case will become a Wave 3 respondent is estimated using logistic regression, which is commonly used to predict response propensities (Groves et al., 2008; Wagner, 2010). Models were fit to the sets of Wave 3 finalized and interim cases, separately for Wave 2 responding adults and youth, to predict the probability of an interim case becoming a respondent as a function of respondent characteristics from earlier waves such as age, sex, race/ethnicity, tobacco use status, education and general health condition (both for continuing adults only), number of

adults in the household, number of contact attempts before a completed interview was obtained, and ever having refused participation at the person level. During the model fitting process, a conservative approach was taken by setting the response variable equal to one for respondents and zero for both finalized nonrespondents and interim cases. When calculating the predicted response rate, the predicted value from the logistic regression model was used for the interim cases, with the exception of interim refusals and persons who are difficult to locate who were treated as finalized nonrespondents.

Results

Tables 3-2 and 3-3 provide predicted retention rates for continuing adults and continuing youth who responded in Wave 2. In addition to the overall row, each table includes rows on tobacco use status, age, sex, race, and ethnicity subgroups based on Wave 2 data for characteristics where information was updated (age and tobacco use status), and Wave 1 data otherwise. Persons with missing values for these characteristics were excluded from the response rate calculation for that characteristic.

The weighted predicted retention rates among Wave 2 respondents are approximately 92 percent for continuing adults and 93 percent for continuing youth. The unweighted predicted retention rates are approximately 91 percent for continuing adults and 93 percent for continuing youth. The predicted retention rate for continuing adults is higher than the projected retention rate of 86 percent provided in the Revision Request to OMB for Wave 3; the predicted retention rate for continuing youth is slightly higher than the projected retention rate of 91 percent. The predicted retention rates are sensitive to the models used for predicting the response propensities among the interim cases, and the estimates of retention rates will be more accurate as more information accrues.

As shown in Tables 3-2 and 3-3, the variability among predicted retention rates for subgroups is small. For continuing adults, females appear to have a slightly higher predicted retention rate than males, and persons of ‘other’ race have a lower predicted retention rate than Whites and Blacks. Current established users of tobacco at Wave 2 also appear to have a lower predicted retention rate. However, these apparent differences depend largely on the disposition of the interim cases and no

Table 3-2. PATH Study Wave 3 predicted retention rates by respondent characteristics: Adult Interview (continuing adults)

Characteristic ^a	A: Adult Interviews completed (n)	B: Interim likely to be completed ^b (n)	C: Finalized nonresponse (n)	D: Interim cases (n)	Unweighted predicted retention rate for Wave 3 ^c (%)	Weighted predicted retention rate for Wave 3 ^c (%)
Overall	4,074	393	179	660	90.9	91.8
Tobacco use status ^d						
Current established user	1,740	153	87	284	89.7	89.9
Not current established user	2,280	236	89	369	91.9	92.4
Age						
18-24	1,182	143	40	248	90.1	90.7
25-44	1,383	143	52	229	91.7	93.4
45-64	1,097	86	45	143	92.0	92.2
65+	411	21	42	40	87.7	88.6
Sex						
Male	1,994	190	93	335	90.2	90.3
Female	2,077	202	86	324	91.6	93.0
Race						
White alone	2,915	272	139	448	91.0	92.1
Black alone or in combination	676	65	24	107	91.8	92.2
Other	360	47	12	89	88.4	88.2
Ethnicity						
Hispanic	722	81	20	137	91.4	93.2
Non-Hispanic	3,293	308	155	519	90.8	91.5

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Adult Extended Interview. The age information and tobacco use status are as reported in the Wave 2 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Interim likely to be completed is the sum of predicted probabilities of an interim case becoming a respondent over all interim cases.

^c Predicted retention rate = (A+B)/(A+C+D).

^d A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 2 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

Table 3-3. PATH Study Wave 3 predicted retention rates by respondent characteristics: Youth Interview (continuing youth)

Characteristic ^a	A: Youth Interviews completed (n)	B: Interim likely to be completed ^b (n)	C: Finalized nonresponse (n)	D: Interim cases (n)	Unweighted predicted retention rate for Wave 3 ^c (%)	Weighted predicted retention rate for Wave 3 ^c (%)
Overall	1,620	153	47	250	92.5	92.5
Tobacco use status ^d						
Ever user	289	36	13	71	87.2	87.5
Never user	1,260	113	31	174	93.7	93.7
Age						
12-13	682	54	20	88	93.2	92.8
14-17	938	99	27	162	92.0	92.3
Sex						
Male	836	67	21	118	92.7	92.8
Female	781	85	26	131	92.3	92.2
Race						
White alone	1,095	87	36	140	93.0	93.0
Black alone or in combination	273	25	7	44	91.8	91.9
Other	176	29	2	43	92.5	92.4
Ethnicity						
Hispanic	441	56	11	95	90.9	91.0
Non-Hispanic	1,148	95	36	152	93.0	93.1

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The age information and tobacco use status are as reported in the Wave 2 Youth Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Interim likely to be completed is the sum of predicted probabilities of an interim case becoming a respondent over all interim cases.

^c Predicted retention rate = (A+B)/(A+C+D).

^d An 'ever user' is someone who has ever used one or more of the following tobacco products covered by the Wave 2 Youth Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times.

definitive conclusions can be made. The predicted retention rates are similar for most of the subgroups of continuing youth; however, the predicted retention rate for ever users of tobacco at Wave 2 appears to be slightly lower than that for never users.

3.1.2 Predicted Recruitment Rates among Wave 2 Respondents for Aged-up Adults and Aged-up Youth

This section reports recruitment rates for aged-up adults who completed the Wave 2 Youth Interview and are eligible for the Adult Interview at Wave 3, and aged-up youth who were

participating shadow youth at Wave 2 and are eligible for the Youth Interview at Wave 3. The Wave 3 Youth Interview is the first interview for responding aged-up youth, and aged-up adult respondents complete the Adult Interview for the first time. Please refer to Table 3-1 (above) for the status of the aged-up adults and aged-up youth from replicate group 1.

Method

The methods described in Section 3.1.1 for estimating the retention rates were also used to estimate the recruitment rates for aged-up adults and aged-up youth. For aged-up adults, the unweighted recruitment rate is calculated as

$$RRU_{AUA} = (C_{AUA} + \sum_{i=1}^{I_{AUA}} \hat{p}_{i,AUA}) / (C_{AUA} + N_{AUA} + I_{AUA}),$$

where

- C_{AUA} = number of Wave 3 completed cases or sufficient partials among persons who completed the Youth Interview at Wave 2 and were administered the Adult Interview at Wave 3;
- N_{AUA} = number of Wave 3 finalized nonrespondents among Wave 2 youth respondents who were age 18 by the age classification date;
- I_{AUA} = number of Wave 3 interim cases among Wave 2 youth respondents who were age 18 by the age classification date; and
- $\hat{p}_{i,AUA}$ = predicted probability of interim aged-up adult i becoming a respondent.

The unweighted recruitment rate for aged-up youth is calculated as

$$RRU_{AUY} = (C_{AUY} + \sum_{i=1}^{I_{AUY}} \hat{p}_{i,AUY}) / (C_{AUY} + N_{AUY} + I_{AUY}),$$

where

- C_{AUY} = number of Wave 3 completed cases or sufficient partials among persons who were participating shadow youth at Wave 2 and were administered the Youth Interview at Wave 3;
- N_{AUY} = number of Wave 3 finalized nonrespondents among Wave 2 participating shadow youth who were age 12 by the age classification date;
- I_{AUY} = number of Wave 3 interim cases among Wave 2 participating shadow youth who were age 12 by the age classification date; and
- $\hat{p}_{i,AUY}$ = predicted probability of interim aged-up youth i becoming a respondent.

The weighted recruitment rates were calculated by substituting the sum of Wave 1 IPS weights (YIPSWT or SIPSWT²⁰) for the counts of finalized cases in each category, and replacing $\sum_{i=1}^{IAUA} \hat{p}_{i,AUA}$ in the RRU_{AUA} formula by $\sum_{i=1}^{IAUA} YIPSWT_i * \hat{p}_{i,AUA}$, and $\sum_{i=1}^{IAUY} \hat{p}_{i,AUY}$ in the RRU_{AUY} formula by $\sum_{i=1}^{IAUY} SIPSWT_i * \hat{p}_{i,AUY}$.

Results

Table 3-4 provides predicted recruitment rates for the Adult Interview for aged-up adults, and Table 3-5 provides predicted recruitment rates for the Youth Interview for aged-up youth. In addition to the overall row, each table includes rows on sex, race, and ethnicity subgroups; Table 3-4 also includes rows on tobacco use status. There are no rows corresponding to age subgroups in Table 3-4 or Table 3-5 because most of the aged-up adults are 18 years old and most of the aged-up youth are 12 years old; there are no rows for tobacco use status in Table 3-5 because no information was collected about the tobacco use of shadow youth at Wave 2. Information from the Wave 1 Youth Interview was used to define the demographic characteristics for the aged-up adults, and information from the Wave 1 Household Screener was used to define the demographic characteristics for the aged-up youth. Information from the Wave 2 Youth Interview was used to define tobacco use status for aged-up adults. Persons with missing values for these characteristics were excluded from the response rate calculation for that characteristic.

The predicted recruitment rate is approximately 94 percent for aged-up adults, which is higher than the projected recruitment rate of 87 percent in the Revision Request to OMB for Wave 3. The predicted recruitment rate for aged-up youth of 88 percent is slightly lower than the projected rate of 89 percent.

Subgroup recruitment rates in Tables 3-4 and 3-5 for aged-up adults and for aged-up youth, respectively, are similar with the exception of sex. Among aged-up adults, females appear to have a higher recruitment rate than males; the reverse pattern is observed among aged-up youth.

²⁰ SIPSWT is the name of the Wave 1 IPS weight for shadow youth. The construction of these weights is not described in the 2015 Interim Report; however, they were created analogously to the youth IPS weights, YIPSWT.

Table 3-4. PATH Study Wave 3 predicted recruitment rates by respondent characteristics: Adult Interview (aged-up adults)

Characteristic ^a	A: Adult Interviews completed (n)	B: Interim likely to be completed ^b (n)	C: Finalized nonresponse (n)	D: Interim cases (n)	Unweighted predicted recruitment rate for Wave 3 ^c (%)	Weighted predicted recruitment rate for Wave 3 ^c (%)
Overall	301	41	8	53	94.4	94.1
Tobacco use status ^d						
Ever user	133	18	1	25	94.8	94.3
Never user	161	22	6	27	94.5	94.3
Sex						
Male	155	20	7	27	92.6	92.0
Female	145	21	1	26	96.3	96.3
Race/ethnicity						
Non-Hispanic White alone	145	25	4	31	94.6	94.7
Other	153	15	4	22	94.1	93.3

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The age information and tobacco use status are as reported in the Wave 2 Youth Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Interim likely to be completed is the sum of predicted probabilities of an interim case becoming a respondent over all interim cases.

^c Predicted recruitment rate = (A+B)/(A+C+D).

^d An 'ever user' is someone who has ever used one or more of the following tobacco products covered by the Wave 2 Youth Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks. A 'never user' is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times.

Table 3-5. PATH Study Wave 3 predicted recruitment rates by respondent characteristics: Youth Interview (aged-up youth)

Characteristic ^a	A: Youth Interviews completed (n)	B: Interim likely to be completed ^b (n)	C: Finalized nonresponse (n)	D: Interim cases (n)	Unweighted predicted recruitment rate for Wave 3 ^c (%)	Weighted predicted recruitment rate for Wave 3 ^c (%)
Overall	311	33	28	52	88.0	88.1
Sex						
Male	174	15	12	23	90.6	90.6
Female	137	18	16	29	85.1	85.2
Race/ethnicity						
Non-Hispanic White alone	148	14	14	21	88.4	88.4
Other	163	19	14	31	87.7	87.8

^a The characteristics are as reported in the Household Screener at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Interim likely to be completes is the sum of predicted probabilities of an interim case becoming a respondent overall all interim cases.

^c Predicted recruitment rate = (A+B)/(A+C+D).

3.1.3 Predicted Retention and Recruitment Rates among Wave 1 Respondents

Sections 3.1.1 and 3.1.2 describe the predicted retention and recruitment rates, respectively, among Wave 2 respondents. Most Wave 2 nonrespondents are also eligible for Wave 3 data collection (see Section 1.2.2), and experience in the field to date suggests that the PATH Study will be successful in obtaining completed interviews from about 10 percent of these cases. Table 3-6 presents the counts of Wave 2 respondents and nonrespondents by participant type at Wave 3 among the replicate group 1 cases released by April 29, 2016. The table also shows the predicted recruitment and retention rates among all Wave 1 respondents, obtained by combining predicted results for Wave 2 respondents with predicted numbers of completed Wave 3 interviews among the Wave 2 nonrespondents. Due to small sample sizes, results for subgroups of Wave 2 nonrespondents among each participant type are not presented.²¹

Table 3-6. PATH Study Wave 3 predicted retention and recruitment rates among Wave 1 respondents

Group	Wave 2 respondents		Wave 2 nonrespondents		Wave 1 respondents
	A: Unweighted count (n)	B: Unweighted predicted response rate (%)	C: Unweighted count (n)	D: Assumed response rate (%)	Unweighted predicted response rate ^a (%)
Continuing adults	4,913	90.9	1,175	10	75
Continuing youth	1,917	92.5	212	10	84
Aged-up adults	362	94.4	137	10	71
Aged-up youth	391	88.0	138	10	68

^a Predicted response rate = (A*B+C*D)/(A+C).

Predicted retention and recruitment rates for Wave 1 respondents in Table 3-6 provide perspective on the Wave 3 data collection efforts. However, these predicted rates depend on the assumed degree of success in obtaining completed Wave 3 interviews from Wave 2 nonrespondents (10 percent), which is based on small sample sizes and may change by the end of the field period. Section 2.1 discusses the Wave 2 response rates; the remainder of Section 3 addresses retention and recruitment from Wave 2 to Wave 3 and therefore focuses on the results for Wave 2 respondents, presented in Sections 3.1.1 and 3.1.2.

²¹ The assumed response rate of 10 percent for the Wave 2 nonrespondents does not have a weighted counterpart; therefore, only unweighted predicted response rates were computed for the Wave 1 respondents.

3.1.4 Biospecimen Collections

This section addresses response rates for the collection of urine and blood specimens from continuing adults and aged-up adults who completed a Wave 3 Adult Interview and were asked to provide a specimen. The PATH Study requests a urine specimen from a subsample of continuing adults in Wave 3 who provided urine specimens at a previous wave; it also requests urine and blood specimens at Wave 3 from all aged-up adults.

Method

The response rates were calculated using the following formula:

$$\text{RRU} = \frac{\text{(Number of adults who provided a specimen)}}{\text{(Number of adults from whom a specimen was requested)}}$$

The urine response rate for continuing adults is based on the 2,140 adults who, as of April 29, 2016, were asked to provide a urine specimen following their Wave 3 Adult Interview. Similarly, the urine and blood response rates for aged-up adults are based on the 326 aged-up adults who completed the Wave 3 Adult Interview as of this date.

Tables 3-7 and 3-8 provide unweighted response rates²² for the biospecimen collections. In addition to the overall row, each table includes rows on tobacco use status, sex, race, and ethnicity subgroups. All persons asked to provide a biospecimen completed the Wave 3 Adult Interview so the subgroup definitions use Wave 3 data for characteristics where information was updated (age and tobacco use status), and Wave 1 data otherwise. Table 3-7 includes rows on age subgroups; this is not necessary for Table 3-8 because the age range among Wave 3 aged-up adults is narrow. Adults with missing values for such characteristics were excluded from the response rate calculation for that characteristic.

²² Weighted response rates are not provided because the subset of continuing adults asked to provide a urine specimen at Wave 3 does not represent a readily interpretable portion of the population. Similarly, while most aged-up adults are age 18 at Wave 3, some are older due to the actual time elapsed between the Wave 2 and Wave 3 interviews, or due to nonresponse at Wave 2.

Table 3-7. PATH Study Wave 3 response rates by respondent characteristics: Urine collection (continuing adults)

Characteristic ^a	A: Adults requested to provide urine (n)	Urine	
		B: Urine collected (n)	Unweighted response rate for Wave 3 ^b (%)
Overall	2,140	2,072	96.8
Tobacco use status ^c			
Current established user	1,158	1,135	98.0
Not current established user	929	888	95.6
Age			
18-24	596	573	96.1
25-44	838	824	98.3
45-64	548	528	96.4
65+	158	147	93.0
Sex			
Male	1,093	1,050	96.1
Female	1,046	1,021	97.6
Race			
White alone	1,514	1,463	96.6
Black alone or in combination	355	348	98.0
Other	203	195	96.1
Ethnicity			
Hispanic	410	391	95.4
Non-Hispanic	1,698	1,650	97.2

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Adult Extended Interview. The age information and tobacco use status are as reported in the Wave 3 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Response rate = B/A.

^c A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 3 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

Table 3-8. PATH Study Wave 3 response rates by respondent characteristics: Biospecimen collections (aged-up adults)

Characteristic ^a	A: Adult Interviews completed (n)	Urine		Blood	
		B: Collected (n)	Unweighted response rate for Wave 3 ^b (%)	B: Collected (n)	Unweighted response rate for Wave 3 ^b (%)
Overall	326	284	87.1	156	47.9
Tobacco use status ^c					
Current established user	47	44	93.6	26	55.3
Not current established user	265	227	85.7	122	46.0
Sex					
Male	167	147	88.0	78	46.7
Female	158	137	86.7	78	49.4
Race					
White alone	212	182	85.8	99	46.7
Black alone or in combination	55	50	90.9	27	49.1
Other	41	34	82.9	19	46.3
Ethnicity					
Hispanic	96	82	85.4	52	54.2
Non-Hispanic	227	200	88.1	101	44.5

^a The sex, race, and ethnicity characteristics are as reported in the Wave 1 Youth Extended Interview. The tobacco use status is as reported in the Wave 3 Adult Interview. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b Response rate = B/A.

^c A tobacco user is defined as someone who uses one or more of the following tobacco products covered by the Wave 3 Adult Interview: cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly.

Results

The projected response rates for biospecimen collections in the Revision Request for Wave 3 were 97 percent for urine collection among continuing adults and, among aged-up adults, 83 percent for the collection of urine and 43 percent for the collection of blood. To date, 97 percent of the continuing adults asked to provide urine specimens have done so, meeting the projected response rate. Among the aged-up adults, the response rates for urine and blood collection are 87 percent and 48 percent, respectively; the response rates to date for both biospecimen collections exceed the projected response rates. Subgroup differences in response rates may be exaggerated by small sample sizes; however, current established tobacco users at Wave 3 appear to have higher response rates to the urine and blood collections. Hispanics appear to be more willing to provide blood specimens than other aged-up adults and Blacks appear to be more willing to provide urine specimens.

3.2 Nonresponse Bias Analysis

The standard approach for an analysis of nonresponse bias in a longitudinal cohort study such as the PATH Study would be to compare Wave 3 respondents with Wave 3 nonrespondents with respect to characteristics from a previous wave (Bose and West, 2002; Javitz and Wagner, 2005; Brownstein et al., 2009). Because some of the cases fielded at Wave 3 were nonrespondents at Wave 2, Wave 1 characteristics are compared. At the mid-point of Wave 3 data collection, there are a number of interim cases yet to be finalized as either respondents or nonrespondents. The number of finalized nonrespondents is small and does not include interim cases that will ultimately be nonrespondents. For this analysis, Wave 3 respondents are compared with the finalized nonrespondents. To explore the sensitivity of results to the disposition of the interim cases, Wave 3 respondents are also compared with *provisional nonrespondents*, defined to be the set of finalized nonrespondents plus interim refusals and persons who are difficult to locate. Some of the interim cases among provisional nonrespondents are expected to complete the Wave 3 interview; however, they are more likely to require intensive contact tracing and follow-up efforts than are other interim cases. For this reason, they are considered to be more similar to finalized nonrespondents than are other interim cases.

Sections 2.2 and 2.3 of the 2015 Interim Report describe the weight construction for Wave 1 of the PATH Study. The final raked weights from Wave 1 were designed to reduce the potential nonresponse bias from Wave 1. For Wave 3, the nonresponse bias analysis uses the Wave 1 final

weights and presents results for demographic and tobacco use subgroups defined using Wave 1 characteristics. This is necessary because Wave 2 final weights and characteristics are not available for Wave 2 nonrespondents. Differences between the weighted estimates of Wave 1 characteristics for Wave 3 respondents and nonrespondents therefore identify characteristics that might be associated with nonresponse bias due to attrition between Wave 1 and Wave 3 of the study, after compensating for Wave 1 nonresponse and possible undercoverage.

The statistical method used for detecting differences between the characteristics of the Wave 3 respondents and nonrespondents is the same as was used for comparing Wave 2 respondents and nonrespondents in Section 2. Point estimates were calculated using the Wave 1 final weights as described above. The corresponding replicate weights were used to calculate variances, and account for the complex sampling features of stratification and clustering. Precisions for the estimates are reported using 95 percent confidence intervals based on the modified Wilson confidence interval approach. SAS software version 9.4 was used to calculate all point estimates and confidence intervals. An estimated difference between Wave 3 respondents and nonrespondents is considered statistically significant if the confidence interval for the difference excludes zero.

Tables 3-9 and 3-10 compare Wave 1 demographic characteristics, health insurance coverage, and tobacco use rates for Wave 3 adult respondents who were adults at Wave 1 with the finalized nonrespondents and with the provisional nonrespondents (who were adults at Wave 1). Tables 3-11 and 3-12 present similar comparisons for the persons who were interviewed as youth at Wave 1 (these cases are a mixture of Wave 3 continuing youth, continuing adults who were Wave 2 aged-up adults, and Wave 3 aged-up adults); the cases are combined for this analysis because the number of finalized nonrespondents among aged-up adults is too small (less than 90) to allow for meaningful comparisons.

Table 3-9. Comparison of Wave 3 Adult Interview respondents with finalized and provisional nonrespondents (Wave 1 adults)

Characteristic at Wave 1 ^a	Wave 3 respondents to Adult Interview		Wave 3 finalized nonrespondents to Adult Interview		Difference in weighted percentages [respondents – finalized nonrespondents] [95% confidence interval]	Wave 3 provisional nonrespondents to Adult Interview		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]		Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Sex								
Male	1,911	47.1% [45.1%, 49.2%]	515	53.6% [49.3%, 57.9%]	-6.4% [-11.3%, -1.6%]	680	54.3% [50.2%, 58.4%]	-7.2% [-11.9%, -2.5%]
Female	1,974	52.9% [50.8%, 54.9%]	416	46.4% [42.1%, 50.7%]	6.4% [1.6%, 11.3%]	522	45.7% [41.6%, 49.8%]	7.2% [2.5%, 11.9%]
Age group								
18-24	1,043	11.7% [10.7%, 12.8%]	240	10.4% [8.6%, 12.5%]	1.3% [-0.5%, 3.1%]	339	11.6% [9.9%, 13.6%]	0.1% [-1.6%, 1.9%]
25-44	1,347	35.0% [33.1%, 36.9%]	295	30.6% [26.9%, 34.6%]	4.4% [0.0%, 8.7%]	396	31.9% [28.6%, 35.5%]	3.0% [-1.1%, 7.1%]
45-64	1,130	35.6% [33.6%, 37.8%]	259	33.3% [29.3%, 37.5%]	2.3% [-2.7%, 7.4%]	315	32.6% [28.8%, 36.5%]	3.1% [-1.6%, 7.8%]
65+	368	17.7% [15.9%, 19.6%]	137	25.7% [22.1%, 29.6%]	-8.0% [-12.4%, -3.6%]	152	23.9% [20.8%, 27.3%]	-6.2% [-10.1%, -2.3%]
Race/ethnicity								
Non-Hispanic White alone	2,328	66.3% [63.4%, 69.1%]	634	73.1% [68.5%, 77.3%]	-6.8% [-11.1%, -2.4%]	781	69.8% [65.2%, 74.0%]	-3.5% [-7.8%, 0.8%]
Other	1,494	33.7% [30.9%, 36.6%]	272	26.9% [22.7%, 31.5%]	6.8% [2.4%, 11.1%]	393	30.2% [26.0%, 34.8%]	3.5% [-0.8%, 7.8%]
Health insurance								
Yes	3,066	85.4% [83.7%, 86.8%]	723	85.7% [82.8%, 88.1%]	-0.3% [-3.1%, 2.5%]	929	85.2% [82.4%, 87.6%]	0.2% [-2.7%, 3.0%]
No	801	14.6% [13.2%, 16.3%]	192	14.3% [11.9%, 17.2%]	0.3% [-2.5%, 3.1%]	253	14.8% [12.4%, 17.6%]	-0.2% [-3.0%, 2.7%]
Education								
< HS or GED	801	17.1% [15.3%, 19.0%]	169	17.9% [14.6%, 21.9%]	-0.9% [-5.3%, 3.6%]	256	19.6% [16.6%, 23.1%]	-2.5% [-6.6%, 1.5%]
HS	826	21.7% [20.1%, 23.5%]	240	26.2% [22.5%, 30.1%]	-4.4% [-8.6%, -0.3%]	296	25.1% [21.7%, 28.8%]	-3.3% [-7.4%, 0.7%]
Some college, no degree	1,391	31.6% [29.6%, 33.6%]	326	31.3% [27.4%, 35.4%]	0.3% [-4.4%, 5.0%]	410	31.6% [27.9%, 35.4%]	0.0% [-4.4%, 4.4%]
Bachelor degree +	859	29.6% [27.2%, 32.1%]	188	24.6% [21.3%, 28.3%]	5.0% [0.7%, 9.3%]	230	23.7% [20.6%, 27.1%]	5.9% [1.8%, 9.9%]

Table 3-9. Comparison of Wave 3 Adult Interview respondents with finalized and provisional nonrespondents (Wave 1 adults) (continued)

Characteristic at Wave 1 ^a	Wave 3 respondents to Adult Interview		Wave 3 finalized nonrespondents to Adult Interview		Difference in weighted percentages [respondents – finalized nonrespondents] [95% confidence interval]	Wave 3 provisional nonrespondents to Adult Interview		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]		Un-weighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Tobacco use status ^b								
Current established use	1,717	22.7% [21.3%, 24.1%]	450	25.4% [22.3%, 28.7%]	-2.7% [-6.2%, 0.8%]	592	27.0% [24.0%, 30.3%]	-4.4% [-7.8%, -0.9%]
Not current established user	2,066	77.3% [75.9%, 78.7%]	449	74.6% [71.3%, 77.7%]	2.7% [-0.8%, 6.2%]	566	73.0% [69.7%, 76.0%]	4.4% [0.9%, 7.8%]

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A ‘current established user’ of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 3-10. Comparison of Wave 1 tobacco use* rates for Wave 3 Adult Interview respondents with finalized and provisional nonrespondents (Wave 1 adults)

Characteristic at Wave 1 ^a	Wave 3 respondents to Adult Interview		Wave 3 finalized nonrespondents to Adult Interview		Difference in weighted percentages [respondents – finalized nonrespondents] [95% confidence interval]	Wave 3 provisional nonrespondents to Adult Interview		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]		Unweighted count	Weighted percentage, using adult Wave 1 final weights [95% confidence interval]	
Overall	3,783	22.7% [21.3%, 24.1%]	899	25.4% [22.3%, 28.7%]	-2.7% [-6.2%, 0.8%]	1,158	27.0% [24.0%, 30.3%]	-4.4% [-7.8%, -0.9%]
Sex								
Male	1,859	27.1% [25.0%, 29.3%]	497	29.2% [24.9%, 34.0%]	-2.2% [-7.5%, 3.2%]	655	31.6% [27.3%, 36.2%]	-4.5% [-9.7%, 0.7%]
Female	1,921	18.7% [17.0%, 20.5%]	401	20.8% [17.1%, 25.2%]	-2.1% [-6.2%, 2.0%]	502	21.6% [17.9%, 25.8%]	-2.9% [-6.9%, 1.1%]
Age group								
18-24	1,026	29.4% [26.1%, 32.9%]	238	39.7% [33.2%, 46.7%]	-10.4% [-18.0%, -2.7%]	335	41.0% [35.2%, 47.0%]	-11.6% [-18.3%, -4.9%]
25-44	1,315	26.0% [23.3%, 28.9%]	289	29.5% [24.3%, 35.4%]	-3.5% [-10.3%, 3.2%]	383	32.3% [27.4%, 37.6%]	-6.3% [-12.5%, -0.0%]
45-64	1,089	23.4% [20.9%, 26.0%]	244	26.8% [21.5%, 32.8%]	-3.4% [-9.6%, 2.7%]	297	27.5% [22.5%, 33.2%]	-4.2% [-10.1%, 1.8%]
65+	353	9.9% [7.2%, 13.5%]	127	12.1% [7.5%, 18.9%]	-2.2% [-7.1%, 2.8%]	142	12.0% [7.7%, 18.4%]	-2.1% [-6.8%, 2.6%]
Race/ethnicity								
Non-Hispanic White alone	2,272	23.1% [21.3%, 25.1%]	613	28.0% [24.1%, 32.3%]	-4.9% [-9.4%, -0.4%]	757	30.2% [26.2%, 34.4%]	-7.0% [-11.5%, -2.6%]
Other	1,452	21.4% [19.1%, 23.9%]	263	18.9% [14.6%, 24.1%]	2.5% [-2.9%, 7.9%]	376	20.4% [16.6%, 24.7%]	1.0% [-3.6%, 5.7%]

^a The characteristics are as reported in the Adult Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

* Defined as current established use of tobacco. A tobacco user is defined as someone who uses one or more of the tobacco products covered by the Wave 1 Adult Extended Interview. A 'current established user' of a given tobacco product is someone who currently uses the product every day or some days and: for cigarettes, has smoked at least 100 cigarettes in their lifetime and, for any other tobacco product, has reported they ever used that product regularly. The products covered by the Wave 1 Adult Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, and dissolvable tobacco.

Table 3-11. Comparison of Wave 3 interview respondents with finalized and provisional nonrespondents (Wave 1 youth)

Characteristic at Wave 1 ^a	Wave 3 respondents		Wave 3 finalized nonrespondents		Difference in weighted percentages [respondents – finalized nonrespondents] [95% confidence interval]	Wave 3 provisional nonrespondents		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]		Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Sex								
Male	1,005	52.1% [49.8%, 54.4%]	117	46.4% [40.2%, 52.8%]	5.7% [-1.1%, 12.6%]	166	47.8% [42.4%, 53.3%]	4.3% [-1.7%, 10.3%]
Female	918	47.9% [45.6%, 50.2%]	134	53.6% [47.2%, 59.8%]	-5.7% [-12.6%, 1.1%]	180	52.2% [46.7%, 57.6%]	-4.3% [-10.3%, 1.7%]
Age group								
12-13	691	35.2% [33.1%, 37.3%]	89	34.5% [28.3%, 41.2%]	0.7% [-6.2%, 7.6%]	122	34.2% [29.1%, 39.6%]	1.0% [-4.9%, 6.8%]
14-17	1,236	64.8% [62.7%, 66.9%]	163	65.5% [58.8%, 71.7%]	-0.7% [-7.6%, 6.2%]	226	65.8% [60.4%, 70.9%]	-1.0% [-6.8%, 4.9%]
Race/ethnicity								
Non-Hispanic White alone	936	55.9% [52.2%, 59.6%]	139	61.4% [53.9%, 68.4%]	-5.4% [-13.4%, 2.5%]	180	58.2% [51.3%, 64.8%]	-2.3% [-9.4%, 4.8%]
Other	949	44.1% [40.4%, 47.8%]	110	38.6% [31.6%, 46.1%]	5.4% [-2.5%, 13.4%]	164	41.8% [35.2%, 48.7%]	2.3% [-4.8%, 9.4%]
Tobacco use status ^b								
Ever user	396	21.0% [18.6%, 23.6%]	56	22.2% [16.4%, 29.3%]	-1.2% [-8.1%, 5.8%]	87	25.3% [20.4%, 30.9%]	-4.3% [-10.2%, 1.6%]
Never user	1,467	79.0% [76.4%, 81.4%]	191	77.8% [70.7%, 83.6%]	1.2% [-5.8%, 8.1%]	254	74.7% [69.1%, 79.6%]	4.3% [-1.6%, 10.2%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

^b An ‘ever user’ is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. A ‘never user’ is someone who has never used any of those tobacco products. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 3-12. Comparison of Wave 1 tobacco use* rates for Wave 3 respondents with finalized and provisional nonrespondents (Wave 1 youth)

Characteristic at Wave 1 ^a	Wave 3 respondents		Wave 3 finalized nonrespondents		Difference in weighted percentages [respondents – finalized nonrespondents] [95% confidence interval]	Wave 3 provisional nonrespondents		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]		Unweighted count	Weighted percentage, using youth Wave 1 final weights [95% confidence interval]	
Overall	1,863	21.0% [18.6%, 23.6%]	247	22.2% [16.4%, 29.3%]	-1.2% [-8.1%, 5.8%]	341	25.3% [20.4%, 30.9%]	-4.3% [-10.2%, 1.6%]
Sex								
Male	967	21.5% [18.6%, 24.8%]	115	26.2% [18.0%, 36.5%]	-4.7% [-14.9%, 5.5%]	163	28.7% [20.8%, 38.1%]	-7.2% [-16.7%, 2.3%]
Female	892	20.4% [17.3%, 23.9%]	131	18.8% [12.3%, 27.5%]	1.6% [-6.6%, 9.9%]	176	22.4% [16.9%, 29.1%]	-2.0% [-8.7%, 4.7%]
Age group								
12-13	653	7.1% [5.1%, 9.9%]	86	8.0% [3.8%, 15.8%]	-0.9% [-7.2%, 5.4%]	119	10.2% [5.8%, 17.3%]	-3.1% [-9.1%, 3.0%]
14-17	1,210	28.3% [25.2%, 31.6%]	160	29.6% [21.3%, 39.6%]	-1.3% [-11.2%, 8.5%]	221	33.2% [26.4%, 40.8%]	-4.9% [-12.9%, 3.0%]
Race/ethnicity								
Non-Hispanic White alone	907	21.1% [17.9%, 24.6%]	136	26.4% [18.2%, 36.8%]	-5.4% [-15.1%, 4.4%]	177	29.6% [22.3%, 38.2%]	-8.5% [-17.2%, 0.1%]
Other	917	21.2% [18.0%, 24.8%]	108	16.1% [10.2%, 24.6%]	5.1% [-3.4%, 13.6%]	160	19.0% [13.5%, 26.1%]	2.2% [-5.6%, 10.0%]

^a The characteristics are as reported in the Youth Extended Interview at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

* Defined as ever use of tobacco. An 'ever user' is someone who has ever used one or more of the tobacco products covered by the Wave 1 Youth Extended Interview. Ever use of a tobacco product is defined as having ever used the product, even one or two times. The products covered by the Wave 1 Youth Extended Interview are cigarettes, traditional cigars, cigarillos, little filtered cigars, pipes, smokeless tobacco, snus, hookah, e-cigarettes, dissolvable tobacco, bidis, and kreteks.

Table 3-13 compares Wave 1 demographic characteristics for Wave 3 youth respondents who were shadow youth at Wave 1 with the provisional nonrespondents (these cases are a mixture of continuing youth who were Wave 2 aged-up youth and Wave 3 aged-up youth). The numbers of finalized nonrespondents are too small (less than 120 overall) to permit comparisons of respondents and finalized nonrespondents for the persons who were Wave 1 shadow youth. Table 3-13 does not include subgroups for tobacco use status because shadow youth did not self-report information in Wave 1.

Among Wave 1 adults, some trends differ for the comparisons of respondents with finalized nonrespondents and with provisional nonrespondents (see Table 3-9). The estimated percentage of adults who are non-Hispanic Whites is lower for respondents than for finalized nonrespondents but the difference is not statistically significant when the respondents and provisional nonrespondents are compared. A similar pattern is observed for the percentage of adults with high school education. The estimated percentage of Wave 1 adults ages 25 to 44 is higher for respondents than for finalized nonrespondents. The estimated percentage of adults with at least a bachelor's degree is higher for respondents than for both nonrespondent groups. The estimated percentages of males and persons age 65 or older are both lower for respondents than for finalized and provisional nonrespondents.

Respondents have lower rates of current established tobacco use compared to finalized and provisional nonrespondents among non-Hispanic White and young adults (ages 18 to 24) at Wave 1 (see Table 3-10). Overall and among Wave 1 adults ages 25 to 44, the respondents have a lower rate of current established tobacco use than the provisional nonrespondents, but there is no statistically significant difference in the rate of current established tobacco use between the respondents and finalized nonrespondents.

The results in this report are based on preliminary data, and may change as more cases are finalized. If the trends seen among provisional nonrespondents continue as the interim cases are finalized, however, the PATH Study may experience attrition patterns that are similar to those in other longitudinal surveys. Thompson (2015) noted that younger persons and persons with lower educational levels are more difficult to retain in longitudinal surveys. Cunradi et al. (2005) and Young et al. (2006) have found that smokers were less likely to be retained in subsequent waves of surveys than nonsmokers.

Among Wave 1 youth and shadow youth, there is no evidence of nonresponse bias (all confidence intervals for estimates of differences between Wave 3 respondents and finalized or provisional nonrespondents include zero, see Tables 3-11 to 3-13).

Table 3-13. Comparison of Wave 3 Youth Interview respondents with provisional nonrespondents (Wave 1 shadow youth)

Characteristic at Wave 1 ^a	Wave 3 respondents to Youth Interview		Wave 3 provisional nonrespondents to Youth Interview		Difference in weighted percentages [respondents – provisional nonrespondents] [95% confidence interval]
	Unweighted count	Weighted percentage, using shadow youth Wave 1 final weights [95% confidence interval]	Unweighted count	Weighted percentage, using shadow youth Wave 1 final weights [95% confidence interval]	
Sex					
Male	362	49.2% [45.5%, 52.9%]	74	47.8% [39.0%, 56.6%]	1.4% [-8.4%, 11.2%]
Female	351	50.8% [47.1%, 54.5%]	86	52.2% [43.4%, 61.0%]	-1.4% [-11.2%, 8.4%]
Age group					
9-10	353	49.9% [46.3%, 53.6%]	83	52.4% [44.7%, 60.0%]	-2.5% [-11.1%, 6.1%]
11	360	50.1% [46.4%, 53.7%]	77	47.6% [40.0%, 55.3%]	2.5% [-6.1%, 11.1%]
Race/ethnicity					
Non-Hispanic White alone	364	55.6% [51.1%, 59.9%]	77	51.6% [43.7%, 59.4%]	4.0% [-4.1%, 12.0%]
Other	348	44.4% [40.1%, 48.9%]	83	48.4% [40.6%, 56.3%]	-4.0% [-12.0%, 4.1%]

^a The characteristics are as reported in the Household Screener at Wave 1. For each characteristic, the sum of the counts in all the categories may not be equal to the count in the overall row due to missing values.

3.3 Statistical Approach for Addressing Nonresponse

Initial weights for Wave 3 respondents will be adjusted to address nonresponse at Wave 3. Similar to the Wave 2 weighting process, the nonresponse adjustment process will be done in two stages. The first stage is to account for the fact that some of the nonrespondents may not have been eligible for Wave 3 but this could not be formally ascertained because, for example, they could not be located at the time of Wave 3 data collection. Doing so is routine and amounts to adjusting the sample weights of all those whose eligibility status is known for Wave 3 (e.g., Wave 3 respondents, Wave 3 refusals, those who died) to account for nonresponse among those whose eligibility status could not be ascertained. After this first stage of nonresponse adjustment is completed, the second stage is carried out only among those respondents and nonrespondents known to be eligible for Wave 3. This is accomplished by simply removing from the weighting process those people who died, were in a correctional facility, or left the country prior to Wave 3.

Weight adjustments will be computed within cells formed from the cross-classification of variables with the potential for reducing nonresponse bias. Such variables from Wave 1 include age, race, ethnicity, sex, employment status, education level, tobacco use status, household composition, census block characteristics, and the type of interview completed (adult, youth, or none for shadow). Wave 2 data available for both Wave 2 respondents and nonrespondents (e.g., Wave 2 paradata) may also be included.²³ A tree-based classification method will be employed to identify cells that distinguish between subgroups with different propensities to respond to Wave 3 the PATH Study (see Roth et al., 2006 and Schouten and deNoij, 2005). SAS macros will then be used to compute and apply the weighting adjustment factors and identify potential sources of concern in the adjustment process, such as small cell sizes and large adjustment factors.

Procedures similar to those described in the memo on Wave 1 biospecimen weighting procedures submitted to OMB (and approved on October 9, 2015) can be used to address nonresponse among adults asked to provide biospecimens at Wave 3.

²³ The handling of Wave 2 nonrespondents in Wave 3 presents a challenge in the development of nonresponse adjustments for Wave 3 weighting. Because the formation of nonresponse adjustment cells requires information for both Wave 3 respondents and nonrespondents, the only information uniformly available for both groups is from Wave 1. Response status in Wave 2 would serve as a key discriminator in terms of differentiating between propensities to respond at Wave 3 but would produce very large adjustment factors, resulting in an inordinate contribution to the variance.

3.4 Summary of Findings

Response Rates

Because the PATH Study Wave 3 data collection is ongoing, response rates²⁴ for Wave 2 respondents were calculated using predicted response propensities for interim cases, as described in Section 3.1. With one exception, the predicted responses rates exceed the projected response rates provided in the Revision Request to OMB for Wave 3. Table 3-14 indicates that the predicted retention rate for continuing adults is higher than the projected rate, and the predicted retention rate for continuing youth is slightly higher than the projected rate. The predicted recruitment rate for aged-up adults is also higher than the projection, and the predicted recruitment rate for aged-up youth is slightly lower than the projected value.

The response rate for each biospecimen collection is calculated as the percentage of persons requested to give a biospecimen who provided it. For both continuing adults and aged-up adults in Wave 3, the response rates for the biospecimen collections meet or exceed those projected.

Table 3-14. Summary of PATH Study predicted response rates for Wave 3

Group	Unweighted predicted response rate	Weighted predicted response rate	Projected response rate*
Continuing adults, Adult Interview	90.9%	91.8%	86%
Continuing youth, Youth Interview	92.5%	92.5%	91%
Aged-up adults, Adult Interview	94.4%	94.1%	87%
Aged-up youth, Youth Interview	88.0%	88.1%	89%
Continuing adults, urine collection	96.8%	-	97%
Aged-up adults, urine collection	87.1%	-	83%
Aged-up adults, blood collection	47.9%	-	43%

* Provided in the Revision Request to OMB for Wave 3 data and biospecimen collections.

The variability among predicted retention rates for subgroups is small (see Tables 3-2 and 3-3). For continuing adults, females appear to have a slightly higher predicted retention rate than males, and persons of ‘other’ race have a lower predicted retention rate than Whites and Blacks. Current established users of tobacco at Wave 2 also appear to have a lower predicted retention rate. The predicted retention rates are similar for most of the subgroups of continuing youth; however, the predicted retention rate for ever users of tobacco at Wave 2 appears to be slightly lower than that

²⁴ Response rates include retention rates for continuing adults and youth, recruitment rates for aged-up adults and youth, and response rates for providing biospecimens.

for never users. Subgroup recruitment rates for aged-up adults and for aged-up youth are similar with the exception of sex: among aged-up adults, females appear to have a higher recruitment rate than males; the reverse pattern is observed among aged-up youth (see Tables 3-4 and 3-5). Response rates for the Wave 3 biospecimen collections meet or exceed the projected rates for all categories (see Tables 3-7 and 3-8).

Nonresponse Bias Analysis

The nonresponse bias analysis found differences between Wave 1 adult respondents and nonrespondents for a few demographic and tobacco use characteristics (see Tables 3-9 and 3-10), but no differences between those groups for Wave 1 youth or shadow youth (see Tables 3-11 to 3-13). For Wave 1 adults, the estimated percentages of males and persons age 65 or older tend to be lower for the Wave 3 respondents than for the Wave 3 finalized and provisional nonrespondents; whereas the estimated percentage of adults with at least a bachelor's degree is higher for respondents. Wave 1 adults ages 25 to 44 are overrepresented and those with a high school education and non-Hispanic Whites are underrepresented among respondents compared to finalized nonrespondents, but the differences between estimates for respondents and provisional nonrespondents are not statistically significant. Among all adults, estimated rates of current established tobacco use at Wave 1 are not significantly different between respondents and finalized nonrespondents, although the provisional nonrespondents exhibit higher tobacco use rates than the respondents (particularly among 18-44 year olds and non-Hispanic Whites).

Statistical Approach for Addressing Nonresponse

For Wave 3, the weights of respondents will be adjusted to account for the nonrespondents by forming weighting adjustment cells using Wave 1 characteristics of respondents and nonrespondents, as well as Wave 1 and Wave 2 paradata that are available for both Wave 3 respondents and nonrespondents. This weighting will compensate for differences between respondents and nonrespondents with respect to sex, age, other demographic and geographic variables, and selected Wave 1 tobacco use measures.

Conclusions and Implications for Study Going Forward

4

This section summarizes the findings presented in this report on the PATH Study's Wave 1, 2, and Wave 3 response rates, nonresponse bias analyses, and approach to addressing nonresponse. Its conclusions are based on the full sample for Waves 1 and 2, and on the data collected from replicate group 1 during the first 6 months (out of 12) of Wave 3. The section closes with a discussion of the implications of the conclusions for the study going forward.

Conclusions

Response rates in Wave 1 for the Household Screener and Adult Extended Interview were lower than projected in the Non-substantive Change Request to OMB for Wave 1 of the PATH Study. However, nonresponse bias analysis found that many characteristics of respondents in Wave 1 aligned with the 1-year estimates from the 2013 ACS. Exceptions were found for single-person households, education, and ethnicity when comparing PATH Study estimates using IPS weights to 1-year 2013 ACS estimates. Estimates of cigarette smoking among adults in Wave 1 were within the range of estimates found by other national health studies. Moreover, when full sample estimates were adjusted for nonresponse using the raked weights, they more closely approximated the ACS estimates, and adult cigarette smoking rates remained essentially the same.

The response rate for the Wave 1 Youth Interview was higher than projected. Nonresponse bias analysis among youth found that many characteristics of respondents were consistent with the 1-year estimates from the 2013 ACS, with the exception of ethnicity. When the full sample estimates were adjusted for nonresponse among youth, they more closely approximated the 2013 ACS estimates, but the ever-tried cigarette smoking rates for youth in Wave 1 remained lower than those found by other national studies.

The response rates for urine and blood collections in Wave 1 were lower than initially projected. Despite this, nonresponse bias analysis found that many of the characteristics of respondents were generally aligned with estimates of these characteristics from the 1-year 2013 ACS. In addition, when the sample estimates were adjusted for nonresponse, they were found to approximate the ACS estimates more closely.

The retention rate for Wave 2 continuing adults was about three percentage points lower than projected in the Revision Request to OMB for Wave 2; however, the retention rate for continuing youth was almost the same as the projection (lower by about 0.5 percentage points). The recruitment rate for aged-up adults was about one percentage point higher than the projected rate, and the recruitment rate for aged-up youth was approximately six percentage points lower than the projected rate. The response rates for the biospecimen collections in Wave 2 were all higher than projected. The largest differential response rates were for the tobacco use status of aged-up adults asked to provide urine and blood specimens (for each biospecimen, about eight percentage points higher for current established tobacco users than for those who were not).

However, nonresponse bias analysis found that many characteristics of Wave 2 respondents aligned with those of Wave 2 nonrespondents. Some exceptions were found when comparing estimates for continuing adults (current established tobacco use was lower overall, and for males, 18-44 year-olds, and non-Hispanic Whites among respondents) and for continuing youth (ever use of tobacco was lower overall, and for females, 14-17 year-olds, and non-Hispanic Whites among respondents). For continuing adults, males, 18-24 year-olds, and those with high school education were underrepresented among respondents; and 45-64 year-olds, persons with health insurance, and those with at least a bachelor's degree were overrepresented among respondents. However, estimates of cigarette smoking among adults in Wave 2 were within the range of estimates found by other national health studies. Moreover, when the estimates of Wave 1 characteristics based on the full Wave 2 sample were adjusted for nonresponse using the Wave 2 final weights, they were almost identical to the estimates based on the Wave 1 sample and Wave 1 final weights. The Wave 2 adult cigarette smoking rates remained essentially the same using the Wave 2 final weights (compared to using the Wave 1 IPS weights), but the ever-tried cigarette smoking rates for youth remained lower than those found by other national studies. Among aged-up adults, current established tobacco users were more likely to provide urine and blood specimens; urine collection rates were very high among continuing adults.

With one exception, the predicted Wave 3 responses rates exceed the projected response rates provided in the Revision Request to OMB for Wave 3. The predicted Wave 3 retention rate for continuing adults who responded at Wave 2 is about five percentage points higher than projected in the Revision Request to OMB for Wave 3; the predicted retention rate for continuing youth who responded at Wave 2 is about 1.5 percentage points higher than the projection. The estimated recruitment rate for aged-up adults who completed a Youth Interview at Wave 2 is seven percentage points higher than the projected rate, and the estimated recruitment rate for aged-up youth who

participated as shadow youth at Wave 2 is one percentage point higher than projected. The Wave 3 response rates for biospecimens also approximately equal or exceed the projected rates.

There is no evidence of nonresponse bias at Wave 3 for the Wave 1 youth or shadow youth. For Wave 1 adults, the estimated percentage of persons with at least a college degree is higher when calculated from the respondents than from the finalized or provisional nonrespondents. The estimated percentages of males and persons age 65 or older tend to be lower for the Wave 3 respondents than for both nonrespondent groups. Current established use of tobacco is significantly lower among respondents than among provisional nonrespondents, particularly for non-Hispanic Whites and adults ages 18-24 at Wave 1 (for whom tobacco use rates are also significantly different when respondents are compared to finalized nonrespondents). However, as noted, these findings are preliminary pending finalization of interim cases and the remainder of data collection in Wave 3.

Implications for the Study Going Forward

Findings on the response rates, nonresponse bias analysis, and approach to addressing nonresponse for Waves 1, 2, and 3 have important implications for the PATH Study. First, the study should continue implementing new approaches to increase response rates for Wave 3 and subsequent waves. The PATH Study is continually seeking ways to boost the response rates. For example, beginning in Wave 2, it enhanced its efforts to communicate by text message and email with participants who indicated they may be contacted in these ways. In each wave, the PATH Study varies the appearance of materials it provides participants to enhance their interest and engagement. In addition, the study provides a certificate of appreciation to participating youth, and it takes extra steps to interview continuing adults who have relocated to group quarters facilities since their initial interviews. Participants continue to access information about the study on the participant pages of the PATH Study website; adult participants may update their contact information on this website, as well as by returning update forms mailed to them and by calling a toll-free telephone number. The study employs special interviewers with skills in refusal conversion and has a cadre of experienced traveling interviewers that help to augment staffing in specific areas. For tracing difficult-to-locate cases, the study has examined the effectiveness of various on-line search services; as needed, it accesses multiple services for a given case. For Wave 4 and potential future waves, the PATH Study plans to further expand its engagement activities to include the use of short videos, to be displayed on field interviewer laptops (e.g., for participants who are reluctant to participate) and on the study website. In these videos and other materials, the study will provide information on the scope of the

PATH Study, to underline the message participants are contributing to a large and important effort that will increase understanding of tobacco use and health.

Second, data analyses will need to consider carefully the achieved sample sizes and adjustments may be necessary, such as combining some subgroups in analyses. Future planning for longitudinal analyses will need to account for the declining rates of cigarette smoking in the U.S. population in general, notably among youth. For the PATH Study, the trends observed among youth mean smaller sample sizes for youth cigarette smokers and therefore less statistical power for examining within-person changes among this group. At the same time, the larger sample size of youth nonsmokers provides more power for examining the initiation of cigarette smoking over time. The impact of the smaller sample size for youth cigarette smokers must also be considered in the context of the recent and rapid increase of youth use of alternative tobacco products such as e-cigarettes and hookah (CDC, 2015d).

Third, on an ongoing and consistent basis, the PATH Study should examine the sample sizes achieved, as well as those projected for future data collection. In this way, it would be possible to detect differential rates of attrition among subgroups early and make extra efforts to retain persons in subgroups of special analytic interest. The Wave 1 shadow sample, which serves as a reservoir for aged-up youth in subsequent waves, will be exhausted after Wave 3. Therefore, with OMB's approval, the PATH Study plans to replenish the sample at Wave 4. The Wave 4 design includes the selection of a new shadow youth sample, as well as new samples of youth and adults to replenish the Wave 1 sample, taking into account the subgroup sample sizes in the continuing cohort.

References

- American Association of Public Opinion Research (AAPOR). (2015). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 8th ed.*
http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions2015_8theditionwithchanges_April2015_logo.pdf.
- Bose, J., and West, J. (2002). Examining additional nonresponse bias introduced through attrition. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 278-283.
- Brick, J.M., Lê, T., and West, J. (2003). Dealing with movers in a longitudinal study of children. In *Statistics Canada Symposium-challenges in survey taking for the next decade*.
- Brownstein, N., Kalsbeek, W.D., Tabor, J., Entzel, P., Daza, E., and Harris, K.M. (2009). *Non-Response in Wave IV of the National Longitudinal Study of Adolescent Health*.
http://www.cpc.unc.edu/projects/addhealth/data/guides/W4_nonresponse.pdf.
- Center for Behavioral Health Statistics and Quality (CBHSQ). (2015a). *National Survey on Drug Use and Health, 2014*. Rockville, MD: Substance Abuse and Mental Health Services Administration,
<http://www.samhsa.gov/samhsa-data-outcomes-quality/major-data-collections/public-use-files-2014-nsduh>
- Center for Behavioral Health Statistics and Quality (CBHSQ). (2015b). *Results from the 2014 National Survey on Drug Use and Health: Detailed Tables. Prevalence Estimates, Standard Errors, P Values, and Sample Sizes*. Rockville, MD: Substance Abuse and Mental Health Services Administration,
<http://www.samhsa.gov/data/sites/default/files/NSDUH-DefTabs2014/NSDUH-DefTabs2014.pdf>.Centers for Disease Control and Prevention (CDC). (2013). Tobacco product use among middle and high school students — United States, 2011 and 2012. *Morbidity and Mortality Weekly Report*, 62, 893–897. Retrieved from
www.cdc.gov/mmwr/preview/mmwrhtml/mm6245a2.htm.
- Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). (2015a). *National Health and Nutrition Examination Survey Data, 2013-2014*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention,
<http://wwwn.cdc.gov/Nchs/Nhanes/Search/DataPage.aspx?Component=Questionnaire&CycleBeginYear=2013>**Error! Hyperlink reference not valid..**
- Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). (2015b). *National Health Interview Survey Data, 2014*. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention,
http://www.cdc.gov/nchs/nhis/nhis_2014_data_release.htm.
- Centers for Disease Control and Prevention (CDC). (2015c). *National Youth Tobacco Survey Data, 2014*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health,
http://www.cdc.gov/tobacco/data_statistics/surveys/nyts.

- Centers for Disease Control and Prevention (CDC). (2015d). *E-cigarette use triples among middle and high school students in just one year*. Press Release (April 16). Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, <http://www.cdc.gov/media/releases/2015/p0416-e-cigarette-use.html>.
- Cunradi, C. B., Moore, R., Killoran, M., and Ames, G. (2005). Survey nonresponse bias among young adults: the role of alcohol, tobacco, and drugs. *Substance Use & Misuse*, 40, 171-185.
- Currihan, D.B., Nyman, A.L., Turner, C.F., and 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 Opinion Quarterly*, 68, 542-564.
- Fowler, F.J., and Stringfellow, V.L. (2001). Learning from experience: Estimating teen use of alcohol, cigarettes, and marijuana from three survey protocols. *Journal of Drug Issues*, 31, 643-664.
- Groves, R.M., Brick, J.M., Couper, M., Kalsbeek, W., Harris-Kojetin, B., Kreuter, F., Pennell, B., Raghunathan, T., Schouten, B., Smith, T., Tourangeau, R., Bowers, A., Jans, M., Kennedy, C., Levenstein, R., Olson, K., Peytcheva, E., Ziniel, S., and Wagner J. (2008). Issues facing the field: Alternative practical measures of representativeness of survey respondent pools. *Survey Practice*, 1(3).
- Javitz, H., and Wagner, M. (2005). *Analysis of potential bias in the Wave 1 and Wave 2 respondents to the National Longitudinal Transition Study-2 (NLTS-2)*. Menlo Park, CA: SRI International, http://www.nlts2.org/studymeth/nlts2_analysis_bias_respondents.pdf.
- Lundström, S., and Särndal, C-E (1999). Calibration as a standard method for treatment of nonresponse. *Journal of Official Statistics*, 15 (2), 305-327.
- Office of Management and Budget. (2006). *Standards and Guidelines for Statistical Surveys*, available at http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf.
- Roth, S., Montaquila, J., and Chapman, C. (2006). Nonresponse bias in the 2005 National Household Education Surveys Program. Technical Report. (NCES 2007-016). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Ryan, H., Troscclair, A., and Gfroerer, J. (2012). Adult current smoking: Differences in definitions and prevalence estimates—NHIS and NSDUH, 2008. *Journal of Environmental and Public Health*, online.
- Särndal, C.-E., and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. Hoboken, NJ: Wiley.
- SAS Institute, Inc. (2013). *SAS® 9.4 Help and Documentation*. Cary, NC: SAS Institute, Inc.
- Schenker, N., and Gentleman, J.F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, 55, 182-186.

- Schouten, B., and de Nooij, G. (2005). *Nonresponse adjustment using classification trees*. Discussion paper 05001, Statistics Netherlands. Available at www.cbs.nl.
- Song, Y. (2013). *Rotation group bias in smoking prevalence estimates using TUS-CPS*. Paper presented at the Federal Committee on Statistical Methodology Research Conference, Washington DC, paper available at http://www.fcsm.gov/13papers/I3_Song_2013FCSM.pdf, and slides available at http://www.copafs.org/UserFiles/file/fcsm/I3_Song_2013FCSM.pdf.
- Thompson, M.E. (2015). Using longitudinal complex survey data. *Annual Review of Statistics and its Application*, 2, 305-320.
- United Nations. (2005). *Designing Household Survey Samples: Practical Guidelines*. United Nations Publication ST/ESA/STAT/SER.F/98, New York: United Nations. Available at <http://unstats.un.org/unsd/demographic/sources/surveys/Handbook23June05.pdf>.
- United States Department of Commerce, Census Bureau. (2012). *National Cancer Institute-sponsored Tobacco Use Supplement to the Current Population Survey (2010-11)*: <http://appliedresearch.cancer.gov/studies/tus-cps/>.
Technical documentation: <http://www.census.gov/cps/methodology/techdocs.html>.
Retrieved from <http://appliedresearch.cancer.gov/studies/tus-cps/results/data1011/table1.html>.
- Wagner, J. (2010). The fraction of missing information as a monitoring tool for survey data quality. *Public Opinion Quarterly*, 74(2):223-243.
- Wilson, E.B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*, 22, 209-212.
- Young, A.F., Powers, J.R., and Bell, S.L. (2006). Attrition in longitudinal studies: Who do you lose? *Australian and New Zealand Journal of Public Health*, 30, 353-361.

Appendix A Cigarette Smoking Questions in the PATH Study and Other Surveys

Table A-1 lists the questions used to ask about current smoking status of adults in the PATH Study and in the surveys used for comparison and describes the populations included in the estimates from those surveys.

Note that although the questions used to define current cigarette smoking are similar among the surveys, small differences could have an effect on the answers given. In the PATH Study, the question used to establish whether a respondent has smoked at least 100 cigarettes in his or her lifetime has closed response categories:

“How many cigarettes have you smoked in your entire life? A pack usually has 20 cigarettes in it.”

For adults:

1. 1 or more puffs but never a whole cigarette;
2. 1 to 10 cigarettes (about ½ pack total);
3. 11 to 20 cigarettes (about ½ pack to 1 pack);
4. 21 to 50 cigarettes (more than 1 pack but less than 3 packs);
5. 51 to 99 (more than 2 ½ packs but less than 5 packs); and
6. 100 or more cigarettes (5 packs or more).

For youth:

1. 1 or more puffs but never a whole cigarette;
2. 1 cigarette;
3. 2 to 10 cigarettes (about ½ pack total);
4. 11 to 20 cigarettes (about ½ pack to 1 pack);
5. 21 to 50 cigarettes (more than 1 pack but less than 3 packs);

6. 51 to 99 (more than 2 ½ packs but less than 5 packs); and
7. 100 or more cigarettes (5 packs or more).

In TUS-CPS, NHIS, and NHANES, however, the question “Have you smoked at least 100 cigarettes in your entire life?” calls for a yes/no response.

The positioning of the questions also differs among the surveys. In the PATH Study, the cigarette smoking questions are near the beginning of the adult questionnaire, and the respondent knows that the questionnaire is about tobacco use behaviors. In TUS-CPS, the smoking questions are near the beginning of the adult questionnaire on tobacco, but the survey is administered as part of the CPS. In NHIS, the smoking questions follow a long series of questions on health problems (breathing problems, diabetes, hernias, hemorrhoids, etc.). These question contexts may be associated with differences in responses.

Table A-2 lists the questions used to define youth cigarette smoking in the PATH Study, NHANES, NSDUH, and NYTS.

Table A-1. Questions used to define adult current cigarette smoking in the PATH Study, TUS-CPS, NHIS, NHANES, and NSDUH

PATH Study	TUS-CPS*	NHIS	NHANES	NSDUH (original definition)	NSDUH (modified definition)**
Question to define current smoking (answers defining current smoking given in parentheses)					
["Have you ever smoked a cigarette, even one or two puffs?" (Wave 1 question, yes) or "In the past 12 months, have you smoked a cigarette, even one or two puffs" (Wave 2 question, yes)] and "Do you now smoke cigarettes every day, some days, or not at all?" (every day or some days) and "How many cigarettes have you smoked in your entire life? A pack usually has 20 cigarettes in it." (100 or more cigarettes (5 packs or more))	"Have you smoked at least 100 cigarettes in your entire life?" (yes) and "Do you now smoke cigarettes every day, some days, or not at all?" (every day or some days)	"Have you smoked at least 100 cigarettes in your ENTIRE LIFE?" (yes) and "Do you NOW smoke cigarettes every day, some days or not at all?" (every day or some days)	"{Have you/Has SP} smoked at least 100 cigarettes in {your/his/her} entire life?" (yes) and "{Do you/Does SP} now smoke cigarettes every day, some days or not at all?" (every day or some days)	"Have you ever smoked part or all of a cigarette?" (yes) and "During the past 30 days, have you smoked part or all of a cigarette?" (yes)	"Have you ever smoked part or all of a cigarette?" (yes) and "During the past 30 days, have you smoked part or all of a cigarette?" (yes) and "Have you smoked at least 100 cigarettes in your entire life?" (yes)
Age range included in estimate					
18+	18+	18+	18+	18+	18+
Exclusions from population					
The Wave 1 target population included only the U.S. civilian, non-institutionalized population. The target population for Wave 2 was the Wave 1 target population residing in the U.S. at Wave 2, except for those who were incarcerated at that time. Thus, it includes Wave 1 respondents who were on active duty or living in a health care institution (e.g., a nursing home) but not those in a correctional facility at Wave 2.	Includes only the U.S. civilian, non-institutionalized population.	Includes only the civilian noninstitutionalized population residing in the U.S. at the time of the interview. Several segments of the population are excluded, such as: persons in long-term care institutions; persons on active duty with the Armed Forces; persons in correctional facilities; and U.S. nationals living in foreign countries.	Includes only the U.S. civilian, non-institutionalized population.	Includes only the U.S. civilian, non-institutionalized population. Excludes homeless persons who do not use shelters, military personnel on active duty, and residents of institutional group quarters, such as jails and hospitals.	Includes only the U.S. civilian, non-institutionalized population. Excludes homeless persons who do not use shelters, military personnel on active duty, and residents of institutional group quarters, such as jails and hospitals.

Table A-1. Questions used to define “current smoking” in the PATH Study, TUS-CPS, NHIS, NHANES, and NSDUH (continued)

PATH Study	TUS-CPS*	NHIS	NHANES	NSDUH (original definition)	NSDUH (modified definition)**
Proxy responses allowed					
No	Yes	Yes, for individuals physically or mentally incapable of responding.	No	No	No

*Proxies are allowed if fourth callback, the person will not return before closeout, or the household is getting irritated. See http://appliedresearch.cancer.gov/studies/tus-cps/surveys/tuscps_english_2010.pdf, p3.

**The modified definition is given in Ryan et al. (2012).

Table A-2. Questions used to define youth cigarette smoking in the PATH Study, NHANES, NSDUH, and NYTS

PATH Study	NHANES	NSDUH	NYTS
Question to define ever tried cigarette smoking (answers defining ever tried cigarette smoking given in parentheses)			
“Have you ever tried cigarette smoking, even one or two puffs?” (Wave 1 question for all youth, and Wave 2 question for aged-up youth, yes) or “In the past 12 months, have you smoked a cigarette, even one or two puffs?” (Wave 2 question for continuing youth, yes)	“About how many cigarettes have you smoked in your entire life?” (1 or more puffs to 100 or more cigarettes)	Have you ever smoked part or all of a cigarette?” (yes)	“Have you ever tried cigarette smoking, even one or two puffs?” (yes)
Questions for determining whether have smoked in past 30 days			
[“Have you ever tried cigarette smoking, even one or two puffs?” (Wave 1 question for all youth, and Wave 2 question for aged-up youth, yes) or “In the past 12 months, have you smoked a cigarette, even one or two puffs?” (Wave 2 question for continuing youth, yes)] and “When was the last time you smoked a cigarette, even one or two puffs?” (earlier today, not today but sometime in the past 7 days, not in the past 7 days but sometime in the past 30 days)	“On how many of the past 30 days did {you/SP} smoke a cigarette?” (1-30)	“Have you ever tried cigarette smoking, even one or two puffs?” (yes) and [“During the past 30 days, have you smoked part or all of a cigarette?” (yes) or “During the past 30 days, that is since [DATEFILL], on how many days did you smoke part or all of a cigarette?” (1-30)].	“During the past 30 days, on how many days did you smoke cigarettes?” (1-30)
Age range included in estimate			
12-17	12-17	12-17	12-17

Table A-2. Questions used to define youth cigarette smoking in the PATH Study, NHANES, NSDUH, and NYTS (continued)

PATH Study	NHANES	NSDUH	NYTS
Exclusions from population			
<p>The Wave 1 target population included only the U.S. civilian, non-institutionalized population.</p> <p>The target population for Wave 2 was the Wave 1 target population residing in the U.S. at Wave 2, except for those who were incarcerated at that time. Thus, it includes Wave 1 respondents who were on active duty or living in a health care institution (e.g., a nursing home) but not those in a correctional facility at Wave 2.</p>	<p>Includes only the U.S. civilian, noninstitutionalized population.</p>	<p>Includes only the U.S. civilian, noninstitutionalized population. Excludes homeless persons who do not use shelters, military personnel on active duty, and residents of institutional group quarters, such as jails and hospitals.</p>	<p>Only includes youth who are public and private school students enrolled in regular middle schools and high schools in grades 6 through 12 in the 50 U.S. States and the District of Columbia. Alternative schools, special education schools, Department of Defense operated schools, vocational schools that serve only pull-out populations, and students enrolled in regular schools unable to complete the questionnaire without special assistance, are excluded.</p>
Other comments			
	<p>Youth with missing values for the response to the question about number of lifetime cigarettes smoked were excluded from the estimates of ever tried cigarette smoking.</p> <p>Youth with missing values for the response to the question about number of cigarettes smoked in the past 30 days were excluded from the estimates of past 30 day cigarette use unless the value was missing because the youth had never smoked a cigarette in his/her lifetime. Youth who had never smoked were treated as having smoked zero cigarettes in the past 30 days.</p>		<p>Self-administered survey in classroom.</p>