Supporting Statement B
for
Population Assessment of Tobacco and Health (PATH) Study (NIDA)
–
Third Wave of Data Collection

June 10, 2015

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
Email: PATHprojectofficer@mail.nih.gov

Table of Contents

LIST OF ATTACHMENTS
(Appearance in Supporting Statement B)

B. Collections of Information Employing Statistical Methods

The following section focuses on a description of the statistical methods planned for Wave 3 of the PATH Study. Section B.1 describes the baseline or Wave 1 target population of the PATH Study as well as the respondent universe and the expected Wave 1, Wave 2, and Wave 3 sample compositions for age, tobacco-use, and race-ethnicity subgroups. It also discusses the PATH Study's sample design and the expected response rates for Wave 3. Section B.2 describes the procedures for collecting PATH Study data. It presents weighting and estimation procedures, with an elaboration of the degree of precision expected for the analyses of various domains of interest. Section B.3 describes procedures for maximizing the participation and retention of the PATH Study respondents. Section B.4 discusses procedures for evaluating the data collection procedures, including a discussion of nonresponse bias. The final section, Section B.5, presents a list of the statistical consultants contributing to the PATH Study.

B.1 Respondent Universe and Sampling Methods

B.1a Target Population

The Wave 1 target population of the PATH Study is the civilian household population 18 years of age or older and youth 12 to 17 years old in the U.S. (the 50 states and the District of Columbia). College students are sampled through their permanent residence rather than at their dormitory. Active-duty members of the military (Army, Navy, Marines, Air Force, and Coast Guard) are excluded, as are all persons living in institutional and non-institutional group quarters other than college dormitories. Spouses and children of active-duty military living off post in the 50 states and D.C. are covered.

B.1b Respondent Universe and Estimated Sample Composition

Estimates of the youth respondent universe and estimated respondent sample sizes for Waves 1 to 3 of the PATH Study are shown in the second row of Table B-1. The estimated respondent universes are based on estimated population counts of persons eligible for the PATH Study from the 2013 American Community Survey (ACS). As shown in Table B-1, the number of completed interviews with youth 12 to 17 years old at Wave 1 is 13,651. After accounting for Wave 1 shadow sample[1] members who have turned 12, youths interviewed at Wave 1 who have become adults, and expected attrition among the remainder of the Wave 1 youth cohort, the estimated number of completed interviews with youth 12 to 17 years old at Wave 3 is 11,548.

Estimates of the PATH Study adult respondent universe are shown in Table B-2, which presents the number of persons by age, tobacco use, and race domains derived from population projections. There are varying definitions of "tobacco user." Table B-2 provides estimated Wave 1 sample sizes for each of three definitions of interest for the PATH Study. The first, called the "wide net" definition, classifies a person as a tobacco user if he or she has smoked a cigarette, cigar, or pipe, or used smokeless tobacco in the last 30 days; and/or has ever used an e-cigarette, snus, dissolvable tobacco, or smoked tobacco in a hookah. This "wide net" is intended to capture adults who have had experience with tobacco products and who may be at risk of progressing to more frequent use. A "current user" of tobacco is anyone who (1) has smoked at least 100 cigarettes in their lifetime and smokes cigarettes every day or some days, and/or (2) smokes cigars/cigarillos/pipe and/or uses smokeless tobacco every day or some days, and/or (3) uses e-cigarettes, hookah tobacco, snus, and/or dissolvable tobacco every day or some days.[2]

---

[1] The "shadow sample" consists of children who are between the ages of 9 and 11 at the household's Wave 1 interview. These children were not interviewed at Wave 1, but they will be enrolled into the youth cohort in a subsequent wave when they turn age 12.

[2] The definition of tobacco use in the 2010-2011 and earlier versions of the Tobacco Use Supplement to the Current Population Survey (TUS-CPS)

Finally, a “current or experimental user” of tobacco is either (a) anyone who is a “current user” or (b) anyone who has used any of these tobacco products in the past 30 days.

The respondent universe counts in the second column of Table B-2 were computed by estimating the total number of wide-net tobacco users and nonusers in the adult civilian population for each age/race domain from Wave 1 of the PATH Study.[3] The number of completed adult interviews at Wave 1 is 32,320, including 9,112 young adults (18 to 24 year olds) and 5,580 Blacks or African Americans (Black/AA).[4] The number of tobacco users is, of course, largest and the number of non-users is smallest under the “wide net” definition whereas the reverse is true under the “current user” definition. At Wave 3, the corresponding estimates are 27,224 completed adult interviews, with 8,163 young adults and 4,758 Blacks or African Americans. These numbers account for both aging of the Wave 1 sample participants and expected attrition. The PATH Study will generate longitudinal data on a range of tobacco use behaviors within the cohort. Pending the availability of these data, the sample sizes presented in Table B-2 for Wave 3 are estimated using the wide-net tobacco use rates calculated from the Wave 1 sample.

---

encompasses items (1) and (2) of the “current user” definition, but not item (3). The 2014-2015 version of the TUS-CPS includes some of item (3) of the PATH “current user” definition as it now asks about e-cigarettes and dissolvables in the same manner that it asks about cigars, smokeless, and regular pipes. This is expected to continue in future versions of the TUS-CPS.

[3] These estimates were calculated using the final adult weights, which were calibrated to population estimates from the 2013 ACS.

[4] Questions in the PATH Study’s instruments that collect data on race or ethnicity are consistent with the most recent revision of the OMB Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting. However, the term “Black/AA” as used here refers to anyone who chooses African American or Black as a race category (irrespective of whether one or more race categories are chosen and irrespective of their reported ethnicity).

Except for the number of youth in the shadow sample, i.e., 9 to 11 year olds selected at Wave 1 for the purpose of replenishing the 12 to 17 year old youth sample in later waves but not for the purpose of interviews, the sample size estimates in Tables B-1 and B-2 apply to the Wave 1, Wave 2, and Wave 3 *completed interviews* (with or without biological specimens for adults). Specific subgroups in these tables represent the major sampling strata used at the person level at Wave 1. Power projections are provided later in Supporting Statement B for subgroups of potential analytic interest.

Table B-1. PATH Study youth and shadow youth respondent universes and estimated sample sizes at Wave 1, Wave 2, and Wave 3

Group	Respondent universe	Wave 1 sample size	Estimated Wave 2 sample size	Estimated Wave 3 sample size
Children 9-11 (shadow sample)	12,273,575	7,207	4,169	1,767
Youth 12-17	24,852,363	13,651	12,560	11,548

Table B-2. PATH Study adult respondent universes, sample sizes at Wave 1, and estimated sample sizes at Waves 2 and 3

Group	Respondent universe under the "wide net" definition of tobacco use	Wave 1 sample size under the "wide net" definition	Wave 1 sample size under the "current user" definition	Wave 1 sample size under the "current or experimental" user definition	Estimated Wave 2 sample size under "wide net" definition	Estimated Wave 3 sample size under "wide net" definition
18-24 Black/AA user	2,746,432	1,306	871	1,032	1,084	934
18-24 Black/AA non-user	2,263,915	510	945	784	583	660
18-24 non-Black/AA user	15,116,956	5,546	3,720	4,198	4,699	4,142
18-24 non-Black/AA non-user	10,601,252	1,750	3,576	3,098	2,088	2,426
25+ Black/AA user	10,583,809	2,559	1,934	2,297	2,334	2,165
25+ Black/AA non-user	14,970,662	1,205	1,830	1,467	1,086	998
25+ non-Black/AA user	60,671,050	13,675	10,233	11,231	12,297	11,269
25+ non-Black/AA non-user	119,737,510	5,769	9,211	8,213	5,116	4,630
All adults	236,691,585	32,320	32,320	32,320	29,287	27,224

B.1c Sample Design

The Wave 1 sample for the PATH Study was selected using a four-stage, stratified probability sample design involving the selection of: (1) 156 primary sampling units (PSUs) consisting of counties or groups of contiguous counties; (2) 6,049 second-stage sampling units (referred to as segments); (3) 166,088 mailing addresses; and (4) 76,526 eligible sampled persons (SPs) within households occupying dwelling units (DUs) at sampled addresses. In addition to the four stages of selection, a two-phase approach was used for the fourth stage of sampling of adults within households. Interviews were attempted with all youth ages 12 to 17 and adults sampled

at Wave 1. In addition, a “shadow sample” of youth ages 9 to 11 was selected for use as a refresher sample for the youth cohort in later waves of the study. The sampling frames and methods used at each stage of selection for the Wave 1 sample are described in Sections B.1c and B.1d of Supporting Statement B for Wave 1.

The PATH Study is now following up the cohort for Wave 2. Wave 3 will be the final wave of data collection under the current contract period; however, additional follow-up waves, potentially with sample refreshment, may be considered at some point, pending the availability of funding. Youth in the Wave 1 shadow sample who are permitted by a parent or guardian to participate in the Study and reach age 12 by Wave 3 will be interviewed for the first time. Similarly, 16 year olds in the youth sample at Wave 1 who reach age 18 by Wave 3 will receive a Wave 1 version of the adult instrument and be asked to provide urine and blood samples for the first time. In addition, persons who participated in the interview at Wave 1 but did not respond at Wave 2 will be contacted for Wave 3, 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. Specific requests to be withdrawn from the PATH Study will also be respected. There are no plans to refresh the sample at Wave 3.

Plans for biospecimen collections at Wave 3 are described in Section B.2d. Urine and blood samples will be requested from persons who have aged up to the adult cohort at Wave 3. Urine samples will also be requested from a subsample of approximately 17,386 adults who initially provided urine at Wave 1 or Wave 2.

The PATH Study's target population at Wave 1 excluded all active-duty members of the military (Army, Navy, Marines, Air Force, and Coast Guard) and all persons living in group quarters other than college dormitories. Some of the Wave 1 sample members will be active duty at Wave 3 and others will have moved into group quarters living arrangements. Apart from the exceptions noted above, all Wave 1 respondents will be retained as members of the PATH Study cohort for Wave 3 and every effort will be made

to obtain interviews (and biospecimens, as applicable) at Wave 3. This may include, for example, waiting for a study participant to return to the household from a short-term group quarters stay before interviewing him/her for Wave 3.

B.1d Estimated Response Rates

For Wave 1, the unweighted response rates for the screener and extended interviews are 54 percent for households, 75 percent for sampled adults, and 78 percent for sampled youth. The overall Wave 1 response rate is 40 percent for adults and 42 percent for youth (i.e., the product of the screener response rate and the person-level response rate). The response rates for Wave 2 are projected from estimates in the Wave 2 Interim Report of the PATH Study (provided in Attachment 21); projected rates for Wave 3 rely on projections from Wave 2 and on information from the Medical Expenditure Panel Surveys (MEPS) and the National Longitudinal Survey of Youth (NLSY).[5] For continuing adults, projected retention rates for the extended interviews

[5] There are few recent in-person longitudinal surveys in the U.S. that are directly comparable to the PATH Study. MEPS, like the PATH Study, conducts in-person interviews with respondents, and provides recent data on retention of adults in a longitudinal study, where the retention rate is the complement of the attrition rate. Kashihara and Ezzati-Rice (2004), adjusting for the fact that MEPS interviews are conducted every six months rather than annually, estimated year-1 retention for the 1999-2000 MEPS at 90% and year-2 retention at 95%. The retention rates for more recent years of MEPS have been in line with these published rates. Conservative values are used for the PATH Study retention rates to account for differences between the PATH Study and MEPS (such as differences in the frequency of visits and in incentive amounts). The 2012-2013 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) is a cross-sectional survey, but the 2001-2002 NESARC had a follow-up wave in 2004-2005 with a retention rate of 86.7 percent (National Institutes of Health, 2010). The PATH Study's projected retention rates are less than or equal to the rates given for other longitudinal surveys (National Research Council, 2014), or for the British Household Panel Survey (Contoyannis et al. 2004; Lynn 2006, Table 67, where the wave 2, wave 3, and wave 4 retention rates are 87 percent, 91 percent, and 96 percent, respectively). The wave 1 retention for youth in the National Longitudinal Survey of Youth 1997 (NLSY, 2014) was 93 percent, with higher rates for subsequent waves.

are 85 percent for Wave 2 and 86 percent for Wave 3. Projected retention rates for continuing youth are higher: 90 percent for Wave 2 and 91 percent for Wave 3. Recruitment rates for aged-up adults (persons previously interviewed as youth and newly eligible for the adult interview) are estimated to be 86 percent for Wave 2 and 87 percent for Wave 3. The estimated recruitment rates for aged-up youth (shadow youth in the previous wave who have attained age 12) are 88 percent for Wave 2 and 89 percent for Wave 3. The overall Wave 3 response rate is projected to be approximately 30 percent for continuing adults and 34 percent for continuing youth (i.e., the product of the Wave 1 response rate and the expected Wave 2 and Wave 3 retention rates among Wave 1 respondents).

In Table B-3, estimated counts of adults providing biospecimens at Wave 3 are based on several assumptions. First, it is assumed that youth who completed the youth interview at Wave 1 and age up to the adult cohort by Wave 3 will be asked to complete the adult interview and to provide urine and blood samples at Wave 3. Second, among this group, it is assumed that the response rates will be 83 percent for urine and 43 percent for blood.[6] The PATH Study plans to continue blood collection among only new adults (youth aging-up to the adult cohort) in Wave 3.

The PATH Study will request urine samples from a subsample of approximately 17,386 continuing adults who initially provided urine at Wave 1 or Wave 2. (The subsample does not include youth who age up to the adult cohort at Wave 3.) It is assumed that 97 percent of these continuing adults will cooperate with the request for urine.[7]

Retention rates for Wave 3 also account for expected mortality between waves, based on 2011 data in Table 1 of http://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_06.pdf (national vital statistics) and 2011 ACS data. Assumptions differ for adults ages 18 to 24 at Wave 1 (100 percent) and adults ages 25 and over at Wave 1 (99.15 percent).

[6] These unweighted response rates are from the Wave 2 Interim Report of the PATH Study (Attachment 21).

[7] The Wave 3 cooperation rate is assumed to be the same as seen to date for Wave 2, from the Wave 2 Interim Report of the PATH Study (Attachment 21).

Table B-3. Estimated number of respondents for Wave 3

Sampling unit	Percentage or estimated response rate	Estimated number
Primary sampling unit (PSU)		156
Area segments/CDSF segments		6,048[†]
Households with persons sampled at Wave 1		53,686
Adult sample (persons ages 18+)		
Number of youths completing Wave 1 interview		13,651
Number from Wave 1 youth sample eligible for Wave 3 adult interview		4,370
Number from Wave 1 youth sample completing Wave 3 adult interview*	77%	3,356
Number of adults completing Wave 1 interview		32,320
Number of adults completing Wave 1 and Wave 3 interviews**	74%	23,868
Number of adults completing Wave 3 interview		27,224
Number of adults providing urine specimen at Wave 3***		13,805
Number of adults providing blood specimen at Wave 3		765
Youth sample (persons ages 12-17)		
Number of youth permitted to participate in Wave 1 shadow sample		7,207
Number from Wave 1 shadow sample eligible for Wave 3 interview		4,925
Number from Wave 1 shadow sample completing youth interview at Wave 3	79%	3,901
Number of youth completing Wave 1 interview		13,651
Number from Wave 1 youth sample eligible for Wave 3 youth interview		9,281
Number of youth completing Wave 1 and Wave 3 interviews	82%	7,648
Number of youth completing Wave 3 interviews		11,548

[†] One originally sampled segment refused to grant access to the PATH Study for data collection purposes.

* The value of 77 percent of Wave 1 youth completing the Wave 3 adult interview assumes that the youth entering the adult sample at Wave 2 have a retention rate of 0.86 x 0.86, and the youth entering the adult sample at Wave 3 have a retention rate of 0.9 x 0.87. Additionally, it is assumed throughout Table B-3 that 5 percent of the persons who did not respond at Wave 2 will respond at Wave 3 (see Watson and Wooden (2006, 2011) and Statistics New Zealand (2011, Table 8), who report returning rates between 5 percent and 20 percent from other longitudinal surveys). These assumptions result in a retention rate from Wave 1 to Wave 3 of 0.5 x (0.86 x 0.86 +0.9 x 0.87 + [0.14 + 0.1]*0.05) = 0.77.

**The value of 74 percent for the percentage of Wave 1 adults completing the Wave 3 interview assumes a retention rate of 0.85 x 0.86 for adults who respond at Wave 2 and a 5 percent “returner” rate for adults who do not respond at Wave 2. These assumptions result in a retention rate from Wave 1 to Wave 3 of 0.85 x 0.86 + 0.15 x 0.05 = 0.74.

***The number of adults providing urine specimens at Wave 3 is calculated by projecting, using the response rates in the previous footnote, the number of adults in selected tobacco use groups at Waves 1 and 2 who will respond at Wave 3, and adding the projected number of responding aged-up adults at Wave 3. It is assumed that 97 percent of adults who provided urine samples in previous waves, and 83 percent of aged-up adults will provide a urine sample at Wave 3. The number of adults providing blood specimens at Wave 3 is calculated as the product of the estimated number of adults age 18 completing a Wave 3 interview and the assumed blood response rate of 43 percent.

B.2 Procedures for the Collection of Information

This section includes a brief overview and description of the PATH Study's data and biospecimen collection plans for Wave 3. It also discusses the Study's efforts to minimize duplication with other data collections and the burden such collections place on participants and the public.

B.2a Overview

Wave 3 will include all Wave 2 respondents as well as Wave 2 nonrespondents, except, as noted in Section B.1c: (1) SPs who specifically requested withdrawal from the Study, (2) firm or hostile refusers, (3) SPs who were unable to complete a Wave 2 interview in English or Spanish, and (4) SPs with a physical or mental disability that prevents participation in the Study. Wave 3 data and biospecimen collection involves three main components. These are: (1) automated audio computer-assisted self-interviewing (ACASI) extended instruments, with separate instruments for youth and adults; (2) an automated computer-assisted personal interviewing (CAPI) parent instrument; and (3) collection of biospecimens from adults. Collection of biospecimens is not a requirement for adult participation; however, completion of an extended interview is required for biospecimen collection. The data collection components and instruments differ for the four sets of study participants in Wave 3: (1) adult SPs, (2) youth SPs and their parents, (3) youth SPs who age up to the adult cohort, and (4) shadow youth who age up to the youth cohort and their parents.

The primary responsibility of the PATH Study field interviewer is to obtain complete and accurate information from the SPs assigned to them. Meeting this responsibility facilitates proper nonresponse analysis. Field interviewers

receive extensive training on procedures for administration of data collection instruments, including techniques to establish rapport and gain cooperation, to explain the Study's importance to the respondent, and to answer respondent questions or address any concerns.

The PATH Study provides home study, in-person, and web-based training to its field interviewers. Newly hired field interviewers participate in a 16-hour home study training that focuses on the laptop computer and other equipment, Study materials, making initial telephone contact with respondents, administrative tasks, and managing workloads and assignments. Newly hired field interviewers without recent experience in household data collection receive an additional six hours of home study training in general interviewing techniques. After the home study training, interviewers participate in six days of in-person, in-depth training on data and biospecimen collection procedures, including: (1) verifying each sampled person's contact information and scheduling in-person interviews; (2) techniques for obtaining consent; (3) conducting the ACASI extended interviews and CAPI parent interview; (4) collecting UPC codes on tobacco products with the UPC code scanner; (5) collecting and shipping urine samples; (6) scheduling blood collections; (7) issuing respondent incentives; and (8) completing administrative procedures, such as data transmission and reporting to a field supervisor. Experienced field interviewers from Wave 2 will participate in a 15-hour web-based training that focuses on new tasks in Wave 3 as well as on tasks that will continue unchanged from Wave 2 to Wave 3. In addition, field interviewers will each receive a field procedures manual that provides details, with reference materials, on verification of contact information, the interviewing process, questionnaire content, and biospecimen collection. Experienced phlebotomists will receive training on PATH Study procedures for visiting the homes of consenting adults to collect blood samples; this training includes the phlebotomist manual on collecting blood as part of the PATH Study.

Quality control procedures used by the PATH Study to ensure that field interviewers are following specified procedures and protocols include in-person observation, in-person or telephone validation, review of audio recordings of interviews, and review of global positioning system (GPS) data

recorded on interviewer laptops during data collection. Field interviewers who show any area of potential weakness are observed in-person at least one time by a field supervisor or home office staff member. In-person observations are typically concentrated in the early weeks of data collection, so that problems can be detected as early as possible; this provides an opportunity for prompt corrective feedback to the individual field interviewer to help improve his/her on-the-job performance.

Brief in-person or telephone validation interviews are conducted with a sample of adult respondents by PATH Study field interviewers, to confirm that an interview was administered or attempted as reported by the field interviewer. Quality control standards used by the PATH Study require validation of at least ten percent of each field interviewer's data collection attempts with adults to ensure compliance with Study procedures. (See Attachment 2 for Validation Form.)

As part of quality control, selected items from the CAPI portions of data collection with adults are audio-recorded (with the consent of respondents) using computer-assisted recorded interviewing (CARI) and reviewed to assess interviewer performance. As needed (e.g., when a respondent refuses audio-recording), quality control interviews will be conducted by telephone or in-person by trained staff; for some non-complete dispositions (e.g., unable to make contact with an SP), trained staff will validate the disposition in person.

Another quality control tool is the use of GPS data recorded on field interviewers' laptops during data collection. The PATH Study uses these data, with case information, to check that the address location at the time of data collection matches the case address. Cases that fail such data quality checks are reported to field management staff for additional review and follow up. Any suspect findings are fully investigated and resolved.

Additionally, throughout the field period, supervisors remain in close contact with field interviewers. Scheduled weekly telephone conferences are held in which non-finalized cases assigned to field interviewers are reviewed to determine the best approach for working and finalizing the cases. As needed,

based on feedback on field interviewer performance, field management staff retrain field interviewers.

Management staff located at both the home office and remote sites have access to a supervisor management system, including automated management and production reports that are used to monitor the data collection effort. Field interviewers are required to transmit data on a daily basis; data are transmitted to a secure server at the home office to update the automated management reports. These reports include weekly reports on progress during the past week as well as on potentially suspicious field interviewer behavior, such as anomalies in the amount of time between interviews, the scheduling of interviews very early in the morning or late in the evening, or the number of interviews conducted per day.

B.2b Extended Interview

There are similarities and differences in the Wave 3 data collection procedures for (1) adult SPs, (2) youth SPs and their parents, (3) youth SPs who age up to the adult cohort, and (4) shadow youth who age up to the youth cohort and their parents. Approximately 3 months in advance of the anniversary of the earliest Wave 2 interview conducted for an SP's household (household anniversary month),[8] the home office will send a letter to adult SPs and parents of youth SPs to remind them of the upcoming follow-up interview. Approximately 1 month in advance of the household anniversary month, the field interviewer will contact an adult in the household by telephone to update contact information on Study participants and arrange a convenient time for the in-person data collection visit(s) at the SP's home (i.e., at the contacted household or, if the SP has moved, at the SP's new location). (See Attachment 2 for Verification Form.)

[8] To increase operational efficiency in Wave 3, all adult SPs and parents of youth SPs in the same household will be contacted on the same schedule. This schedule is based on the anniversary of the earliest Wave 2 interview conducted with an SP in a household.

Adult SPs

For adult SPs from Wave 1 or Wave 2, at the time of the in-person visit for Wave 3, the field interviewer will: (1) review the main elements of the informed consent, obtained at Wave 1 or Wave 2;[9] (2) administer the Wave 3 adult extended interview, including updating contact information; (3) review the main elements of consent for the biospecimen collection and request a urine sample from a subsample of approximately 17,386 adults who provided urine at Wave 1 or Wave 2; and (4) pay the incentive(s) to the respondent. (The biospecimen collection is discussed further in Section B.2d.) If an adult SP is unavailable or unable to complete the interview at the scheduled time, the field interviewer will attempt to schedule an appointment for a return visit or, at a minimum, determine the best time for a return visit.

After reviewing the main elements of consent, the field interviewer will launch the ACASI extended interview, which begins with an optional tutorial on using ACASI. As required throughout the interview, the field interviewer will remain available to aid the SP in use of ACASI and to respond to questions the SP may have. At the end of the extended interview, the field interviewer will update the SP's contact information.

The adult SP who completes the extended interview will receive $35 (the adult extended interview incentive) on a PATH Study debit card as a thank you for completing the interview as well as a thank you card (Attachment 9). A refusal conversion letter will be sent to adult SPs who initially decline to participate or are difficult to contact (Attachment 19). An adult respondent may also receive $5 for updating his/her contact information on up to two occasions during the year, for a total of $10.

---

[9] The consent documents are framed in terms of the baseline wave. This is appropriate for the two audiences that will be exposed to the documents at Wave 3: (1) respondents who may wish to review the consent document they signed at the first wave in which they participated (Wave 1 or Wave 2), and (2) aged-up participants for whom Wave 3 is their first adult interview.

Youth SPs

For youth SPs who completed the in-person interview at Wave 1 or Wave 2, the field interviewer will: (1) review with the parent the main elements of parent permission for the youth's participation obtained at Wave 1 or Wave 2; (2) review with the parent the main elements of consent for the parent interview, also obtained at Wave 1 or Wave 2; and (3) administer the CAPI parent interview, which includes updating the parent's contact information.[10] Field interviewers cannot conduct youth interviews before completing review of parental permission with the parent. The parent who completes a parent interview for the youth will receive $10 on a PATH Study debit card as a thank you for completing the interview.

If a youth SP with parental permission is available and has time at the visit to complete the interview, the field interviewer will review the main elements of assent with the youth SP, obtained at Wave 1 or Wave 2. The interviewer will then launch the ACASI extended interview. The youth SP who completes the Wave 3 extended interview will receive $25 (the youth extended interview incentive) on a PATH Study debit card and a certificate of appreciation that acknowledges the youth's contributions/community service as a participant in the PATH Study (Attachment 9). The parents of youth SPs will also receive a thank you card. A refusal conversion letter will be sent to parents who are difficult to contact (Attachment 19). A youth SP may also receive $5 on up to two occasions when his/her parent updates the youth's contact information during the year, for a total of $10.

Youth SPs Who Age up to the Adult Cohort

At the in-person visit, the field interviewer will: (1) obtain informed consent (Attachment 12); (2) administer the adult extended interview, including gathering additional contact information about the adult; (3) obtain consent

[10] At Wave 3, if the youth SP is living with a parent other than the one who originally permitted youth participation, the field interviewer obtains permission and consent from the "new" parent and conducts the parent interview with him/her.

for the biospecimen collection; (4) collect the urine sample; (5) arrange a follow-up appointment for a phlebotomist to collect a blood sample; and (6) pay the incentive to the respondent at the completion of the first home visit. (The biospecimen collection is discussed further in Section B.2d.) The incentive for these adults is the same as for continuing adult SPs (i.e., $35 for completing the adult extended interview). The procedures for receiving a thank you card (Attachment 9), refusal conversion letter (Attachment 19), and incentive for updating contact information are the same for youth SPs who age up to the adult cohort and continuing adult SPs.

Shadow Youth Who Age up to the Youth Cohort

The procedures for shadow youth who age up to the youth cohort at Wave 3 are similar to those for continuing youth SPs, except that parental permission, parent consent, and youth assent must be obtained for the first time rather than reviewed from the preceding waves (Wave 1 or Wave 2). The incentive for the parents of youth who age up to the youth cohort is the same as for the parents of continuing youth SPs (i.e., $10 for completing the parent interview).

For the selected youth who has aged up to the youth cohort, following parental permission, the interviewer will obtain youth assent (Attachment 12) and administer the automated ACASI extended youth instrument. The incentive for these youth is the same as for continuing youth SPs (i.e., $25 for completing the youth interview); the aged-up youth also will receive a certificate of appreciation (Attachment 9). The procedures for parents receiving a thank you card, refusal conversion letter, and incentive for updating contact information are the same for youth who have aged up to the youth cohort and for continuing youth SPs.

B.2c Burden Reduction by Avoiding Redundant Data Collection

Wave 3 interviews for continuing adults and youth are designed to build on information collected at Wave 1 and Wave 2. Thus, stable information such as demographic characteristics (e.g., sex and race) is collected only at the respondent's "baseline" or first wave. Similarly, information on lifetime use of tobacco products is not asked again for products a respondent reported having used at his or her baseline wave. Not repeating questions from the previous interview about established characteristics or past behaviors will help keep respondent burden at Wave 3 to an average of 1 hour for the adult interview and 35 minutes for the youth interview.

The parent interview collects information about the parent of a sampled youth, the household, and the youth, as well as contact information to reach the parent for future data collection activities. Because more than one youth may be sampled per household, one parent may be asked to respond to a parent interview for more than one youth. In such cases, the parent will not be asked to provide the same information again, but only information relevant to the particular youth.

B.2d Biospecimens

As noted in Section B.1d, field interviewers will request urine samples at Wave 3 from a subsample of approximately 17,386 adults who initially provided urine at Wave 1 or Wave 2. (See Table B-4.) They will also collect urine and arrange for collection of blood samples from youth SPs who age up to the adult cohort at Wave 3 and consent to biospecimen collection.

Although completion of the extended interview is required from all respondents who choose to participate in the longitudinal cohort, providing biospecimens is voluntary and not a condition of participation. Respondents will receive $25 on a PATH Study debit card for participating in the urine sample component of the Study and $25 for participating in the blood sample component.

Table B-4. Summary of plans for requesting biospecimen collection at Wave 3

Type of respondent	Biospecimen	
	Urine	Blood
Adults who initially provided urine at Wave 1 or Wave 2	Yes, for subsample of approximately 17,386 adults	No
Youth SPs who age up to adult cohort at Wave 3	Yes	Yes

B.2e Weighting and Estimation Procedures

Sample weights will be developed for the PATH Study respondents at each wave to permit estimation for and inference about the population from which the sample was drawn. For Wave 3 of the PATH Study, because no sample replenishment is done, the population of inference is the same as that for the Wave 1 sample after removing deceased persons and other persons ineligible for Wave 3. The sample weights for Wave 3 will be produced to accomplish the following objectives:

- Permit the appropriate development of estimates, taking account of the fact that not all persons in the target population had the same probability of selection in Wave 1;
- Limit the potential for biases arising from differences between cooperating and non-cooperating SPs and households;
- Reduce the variation of the weights and prevent a small number of observations from dominating domain estimates; and
- Facilitate sampling error estimation appropriate to the complex sample design.

The data used in weighting will undergo careful edit, frequency, and consistency checks to prevent errors in the sample weights. The checks will be performed on items to be used in the weighting procedures and will be limited to records that require weights. These checks are important because errors in the weights can affect the PATH Study's estimates.The process for

computing (cross-sectional) Wave 1 weights was described in detail in Section B.2e of Supporting Statement B for Wave 1. The basic steps included:

1. Creating household base weights that are the inverses of the household selection probabilities;

2. Creating household nonresponse-adjusted weights by inflating the household base weights of responding households to compensate for nonresponding households, and "raking" the nonresponse-adjusted weights to population control totals;

3. Creating person base weights by modifying the household nonresponse-adjusted weights to compensate for unequal selection probabilities of SPs;

4. Creating person nonresponse-adjusted weights by inflating the person base weights of responding persons to compensate for nonresponding persons;

5. Creating trimmed weights to reduce any excessive variation in the person nonresponse-adjusted weights;

6. Creating final weights by raking the trimmed weights to population control totals to account for undercoverage and other sources of bias that may remain after applying the above steps; and

7. Creating replicate weights using the balanced repeated replication method for use in variance estimation.

For the PATH Study, one set of Wave 1 weights was created for all youth who completed the Wave 1 interview and another set was created for all adults who completed the Wave 1 interview. Weights were also created for the 9 to 11 year olds selected as part of the shadow sample in Wave 1. Their final Wave 1 weights will serve as the "base weights" for the shadow sample members when they become 12 year olds and join the youth cohort.

The final Wave 1 raked adult, youth, and shadow youth weights will serve as the starting point for the development of weights for Wave 3.[11] Those

[11] Wave 1 weights, rather than Wave 2 weights, are used as the starting point because some persons will provide interviews in Waves 1 and 3 but not in Wave 2.

weights sum to the sizes of the eligible populations of adults, youth, and shadow youth, respectively, for the PATH Study at Wave 1. No sample replenishment is done for Wave 3, so the population of inference is the same as for the sample at Wave 1.

The second step in creating the Wave 3 weights will be to adjust the weights of the respondents to account for attrition between Wave 1 and Wave 3, which is the standard approach used to compensate for wave nonresponse (see, for example, Kalton, 1986 and Brick, 2013). Särndal and Swensson (1987) discuss approaching nonresponse adjustments as analogous to two-phase sampling, which would allow use of Wave 1 (and possibly Wave 2) interview data from all Wave 1 adults and youth to be used to construct weights for continuing adults and youth who participate at Wave 3. Nonresponse adjustment cells will be formed using variables from Waves 1 and 2, including age, race, ethnicity, sex, employment status, education level, tobacco use, household composition, and census block characteristics. For Wave 1 shadow youth who have their first youth interview at Wave 3, information from the Wave 1 household screener will be used to form nonresponse adjustment cells; for Wave 1 shadow youth who were interviewed as youth at Wave 2, the Wave 2 information will be used to form the nonresponse adjustment cells.

Weight adjustments will be computed within cells formed from the cross-classification of available variables. Tree-based classification software will be employed to identify cells that distinguish between subgroups with different propensities to respond to the PATH Study at Wave 3 (see Roth et al., 2006 and Schouten and deNooij, 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. When necessary, imputation may be used for missing values of variables used in the weighting adjustment process (Judkins et al., 2007).

Analyses will be conducted to assess the sensitivity of estimates to the choice of variables used in the weighting procedure. Micklewright et al. (2012) describe methods that may be used to assess sensitivity to

nonresponse adjustments when large amounts of administrative data are available for the respondents and nonrespondents. These methods can be adapted for studying possible bias from attrition in the PATH Study, where interview data are available for the Wave 1 respondents.

B.2f Expected Levels of Precision of the PATH Study

The PATH Study is designed to produce reliable estimates of between-person differences and within-person changes in tobacco-related attitudes, behaviors, and health conditions among various population subgroups and over time. Many characteristics of interest are dichotomous, having “yes” or “no” outcomes. The percentage of “yes” responses is denoted by p and represents the prevalence estimate for a particular characteristic (e.g., cigarette smoking). Past and current research suggests that most of the characteristics measured in the PATH Study will have magnitudes of prevalence that exceed 10 percent, but there will be a few, such as initiation of tobacco use, that will fall between 1 and 5 percent.

One measure of the precision associated with cross-sectional prevalence rates is relative standard error (RSE), defined as the standard error divided by the prevalence estimate and expressed as a percentage. More specifically, RSE(p) = 100 * Standard Error (p)/p, where the standard error is given by the square root of the variance of the estimate, taking into account the complex sample design of the PATH Study. A measure of power associated with longitudinal analyses of change in prevalence rates is the minimum detectable absolute difference (MDAD; see Lipsey, 1990). Herein, the MDADs represent the smallest change (up or down) from a given Wave 1 prevalence rate that can be detected with 80 percent power using a two-sided test for equality of proportions at the 5 percent level of significance, taking into account the Study’s complex sample design. The impact of the various complex features of the sample design on variances, and therefore on RSEs and MDADs, is reflected through inflation factors called design effects (DEFFs). The extent to which these design effects exceed one indicates the extent to which the variance of an estimate based on the

complex sample design is greater than the corresponding variance based on a simple random sample (SRS) design.

Several key features of the PATH Study sampling design contribute to the overall design effect. One of these is clustering at both the PSU and segment levels. In general, for a fixed sample size, the greater the number of units to be sampled per cluster, and the more homogeneous the sampling units are with respect to a characteristic of interest within clusters, the greater the DEFF and hence the inflation in the variance (resulting in decreased precision). The level of homogeneity within a cluster is reflected through two types of intraclass correlations: ρ_1 for PSUs and ρ_2 for segments. Note that ρ_1 and ρ_2 will vary in value for different characteristics of interest. The expected standard errors for prevalence estimates for the PATH Study have been calculated taking into account the contributions due to clustering at both the PSU and segment levels under the assumptions that the intraclass correlations (ρ_1, ρ_2) are (.01, .05). These values were based on estimates taken from various sources in the survey research literature (see, for example, Guilliford et al. [1999] and Thompson et al. [2012]). The calculations reflect that “certainty PSUs” are in fact strata, not PSUs; therefore, there is no contribution to the variance from clustering at the PSU level for these PSUs. Thirty-five of the 156 PSUs selected are certainties, representing 24 percent of the U.S. population.

Another feature is the Wave 1 sampling of adults with different selection probabilities according to their age, race, and tobacco use (as reported by the household screener respondent and as self-reported by the adult at the second phase of screening). The unequal weighting DEFFs due to this feature of the sample design are expected to range from 1.00 to 1.67, depending on the demographic or tobacco use domain of interest. For analyses that combine all adult respondents, this component of the unequal weighting DEFF is approximately 1.81.

A third feature is the restriction that no more than two adults be sampled from a participating household. This requirement contributes to the variability of weights because adults in some multi-person households are

sampled at lower rates than persons of the same age, race, and tobacco use group in single- or two-person households. The unequal weighting DEFFs due to this feature of the sample design are expected to range from 1.00 to 1.02, depending on the domain of interest. For analyses that combine all adult respondents, this component of the unequal weighting DEFF is expected to be negligible (i.e., approximately equal to 1). Note that for analyses of subgroups of race, say by age or sex, these DEFFs will diminish, because generally fewer members of the subgroups will contribute to the clustering effect.

Estimates of precision and power at Wave 3 are calculated after taking into account the DEFFs resulting from the three previously-described sample design features. These estimates are shown in Tables B-5 and B-6, for adults and youth, respectively. The projected RSEs are for a generic statistic estimating a prevalence rate of 15 percent (such as the percentage of the adult population who are every day cigarette smokers). The MDADs are for a generic statistic estimating change from a Wave 1 prevalence rate of 10 percent (such as any non-cigarette tobacco use). Both the RSEs and MDADs presented here are for illustrative purposes.

In Tables B-5 and B-6, the RSEs are for cross-sectional estimates at Wave 3 and the MDADs are for a change from Wave 1 to Wave 3. The subgroups of interest are defined in terms of tobacco-related behaviors, which are subject to change over time. This presents a challenge when trying to estimate the subgroup sample sizes in future waves of the PATH Study, particularly given the recent expansion of tobacco products on the market. Over time, participants sampled as youth will become young adults and those sampled as young adults (18 to 24 years of age) will age into the older group. As a result, variation in weights among members of most subgroups will increase, making it necessary to inflate the assumed values of the DEFFs that are due to unequal weighting. It is not possible to predict the precise inflation factor for each subgroup given the complication of unknown, future rates of quitting, initiating, switching, substituting, or multiple tobacco product use. Inflation factors were computed separately for adults 18 to 24 years old (where at Wave 3, the persons ages 18 and 19 were originally sampled as youth) and for adults ages 25 and over. For adults 18 to 24 years old, the

inflation factors range from 1.01 to 1.33, and for adults ages 25 and over, the inflation factors range from 1.01 to 1.06. The estimates of cross-sectional precision and of detectable changes across waves are presented for a small number of subgroups (i.e., those

Table B-5. Adult sample sizes, relative standard errors (RSEs), and minimum detectable absolute differences (MDADs) at Wave 3*

Group	Wave 3 sample size	RSE on 15% item	MDAD on 10% item
All adults	**27,224**	**2.8**	**0.7**
Wide-net users	19,446	2.7	0.7
Current and experimental users	15,800	2.9	0.7
Menthol cigarette smokers	3,950	4.5	1.1
Dual (smokers and smokeless tobacco users)	950	8.5	2.0
Daily users	9,451	3.3	0.8
Less-than-daily users	4,665	4.2	1.0
Current non-users under wide-net definition	7,778	3.9	0.9
Urine sample providers	13,805	3.4	0.8
Adults ages 18-24	**8,163**	**3.6**	**1.2**

*As indicated in the footnotes to Table B-3, 74 percent of Wave 1 adult respondents are expected to complete the Wave 3 interview.

Table B-6. Youth sample sizes, relative standard errors (RSEs), and minimum detectable absolute differences (MDADs) at Wave 3*

Group	Wave 3 sample size	RSE on 15% item	MDAD on 10% item
All youth	**11,548**	**3.0**	**1.0**
Current tobacco users	981	7.9	2.5
Current cigarette smokers	541	10.5	3.3
Menthol cigarette smokers	327	13.4	4.2
Experimenters	1,281	7.0	2.2
Never cigarette smokers	9,983	3.1	1.0
Susceptible never cigarette smokers	2,084	5.6	1.8
Never tobacco users	6,610	3.6	1.1
Youth ages 12 to 13	**3,901**	**4.3**	**2.1**
Current tobacco users	66	29.7	14.1
Current cigarette smokers	40	38.0	18.1
Menthol cigarette smokers	22	51.6	24.5
Experimenters	170	18.5	8.8
Never cigarette smokers	3,721	4.4	2.1
Susceptible never cigarette smokers	534	10.6	5.1
Never tobacco users	2,435	5.3	2.5
Youth ages 14 to 17	**7,648**	**3.4**	**1.3**

Current tobacco users	915	8.2	3.0
Current cigarette smokers	501	10.9	3.9
Menthol cigarette smokers	305	13.9	5.0
Experimenters	1,111	7.5	2.7
Never cigarette smokers	6,262	3.6	1.3
Susceptible never cigarette smokers	1,550	6.4	2.3
Never tobacco users	4,176	4.2	1.6

*As indicated in Table B-3, the 77 percent retention/response rate is the percentage of youth Wave 1 participants completing the Wave 3 interview.

for which the estimates are expected to be fairly robust to the assumptions made). As a consequence, the estimates herein should be interpreted with caution.

The sample of adult tobacco users at Wave 3 will be sufficiently large to allow analyses for many user subgroups, as well as for persons who are considered at risk for becoming tobacco users. Table B-5 highlights subgroups of potential analytic interest by breaking out sample sizes and measures of precision and power for tobacco users, menthol cigarette smokers, users of both smoked and smokeless tobacco, daily/non-daily tobacco users, and young adults (18 to 24 year olds). In addition, the RSE and MDAD are shown for the sample of adults expected to provide urine specimens at Wave 3. The subgroup sample sizes for the different categories of tobacco users were estimated using data from the full Wave 1 sample of the PATH Study. The “current user” definition from Section B.1b was applied in estimating sample sizes for menthol cigarette smoker, dual and daily user, and less-than-daily user groups; the “wide net” definition was used to estimate the sample sizes for nonuser groups. These are the definitions that give the smallest sample size, and hence the largest RSEs and MDADs, for each of these groups. The estimated RSEs and MDADs for another definition of tobacco user will be smaller than those displayed in the tables. For both RSEs and the MDADs, smaller is better. The RSEs for a 15 percent prevalence rate are at or below 5 percent for most subgroups shown. The MDADs for a 10 percent Wave 1 prevalence rate are mostly at or below 1 percentage point indicating that a two-year change of one percentage point or less can be reliably detected for the subgroups shown. With the exception of the group who provide urine specimens, the adult sample sizes considered in

this section are based on estimates for completed Wave 3 interviews; therefore, the estimates of precision and power apply to projected estimates of tobacco and health outcomes collected with the Wave 3 instruments. As is the case for all the estimates presented in this section, it is expected that precision and power will be reduced for finer divisions of the subgroups (e.g., by gender).

The initial sample of 13,651 youth at Wave 1 was intended both to replenish the adult cohort in future waves of the PATH Study and to provide sufficient power for analyses of youth subgroups. Table B-6 shows Wave 3 sample sizes and measures of precision and power for the youth sample overall and by subgroups of possible interest: tobacco users, cigarette smokers, menthol cigarette smokers, “experimenters,” never cigarette smokers, susceptible never cigarette smokers, and never users of tobacco; the same statistics are shown for each of these subgroups among 12 to 13 year olds and among 14 to 17 year olds. Subgroup sample sizes were estimated using data from the full Wave 1 sample of the PATH Study. For youth, current smokers were defined as youth who have smoked a cigarette within the last 30 days, and current users were youth who have used any tobacco product within the last 30 days. Experimenters were defined as youth who have ever smoked any cigarette, even one or two puffs, but fewer than 100 cigarettes. Susceptibility to initiate cigarette smoking among never smokers was defined as providing any response other than "definitely not" to at least one of the questions: "Do you think that you will try a cigarette soon?", "Do you think you will smoke a cigarette anytime during the next year?" or "If one of your best friends offered you a cigarette, would you smoke it?"

Many of the subgroups are large enough to produce stable estimates. For example, there are approximately 6,610 never users for whom tobacco use initiation rates will be tracked. Tobacco cessation is more of an issue in the older adolescent group (14 to 17 year olds) because there are more tobacco users in that age group than among youth ages 12 to 13, and there are about 915 tobacco users and 501 cigarette smokers whose quitting behavior over time will be monitored. The smallest subgroup presented in Table B-6 that may be of interest is menthol cigarette smokers. Subgroups within the PATH Study are generally expected to provide sufficient precision for

studying tobacco users/experimenters and nonusers among the 12 to 13 year olds; however, subgroups with very small sample sizes, such as 12 to 13 year old menthol cigarette smokers, will likely be combined (e.g., 12 to 17 year old menthol cigarette smokers) to produce estimates with higher precision by type of tobacco use.

The RSEs for a 15 percent prevalence rate among youth 12 to 17 years–old are below 13 percent for most subgroups, and at or below 7 percent for half of them. Among all youth 12 to 17 years old, the sample size overall and in each of the subgroups except current and menthol cigarette smokers is sufficient to detect a two-year change of 3.0 percentage points in a 10 percent Wave 1 behavior overall. This is an important threshold because measures of quitting, initiation, and non-cigarette tobacco use tend to be in this 10 percent range (depending on the definitions used).

B.3 Methods to Maximize Response Rates and Deal with Nonresponse

For Wave 3 and for potential follow-up waves thereafter, the PATH Study will continue to maintain contact with respondents and maximize their retention in the study. The methods currently used by the PATH Study include: (1) tracking participants (through requests to update contact information by visiting the Study website, calling a toll-free number, or sending updated information via mail) and tracing them, as needed; (2) maintaining a sufficiently large field interviewer workforce located in or near the selected PSUs; (3) implementing robust interviewer training and quality control procedures; (4) interviewing in Spanish as well as in English; (5) communicating with participants by mail, telephone, email, and text messaging[12] in advance of in-home data and biospecimen collection visits; (6) continuing to emphasize the importance of biospecimen collections to field interviewers and respondents; (7) as appropriate, interviewing adults

[12]Email or text messaging will only be used with adults who agreed to be contacted this way.

who have relocated to group quarters facilities since their initial interviews; and (8) attempting to convert refusals from adults.

Specific to the PATH Study, OMB's terms of clearance in approving the Wave 2 data collection require NIDA and FDA to report to OMB regarding: "a) the response rates associated with the full baseline wave, including screening, interview completion, and bio-specimen response; b) Wave 2 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." Section B.4 summarizes the results of the PATH Study's Interim Report on these topics.

B.3a Maintaining Participant Engagement and Tracking

The PATH Study seeks to maintain respondent engagement as well as track respondents so they can be contacted for follow-up data and biospecimen collection. The following activities are planned for Wave 3 to maintain respondent engagement:

- Mail a thank you card after the Wave 2 interview and a birthday card to adult SPs and the parents of youth SPs for the youth's birthday;
- In households with only shadow youth, mail a follow-up letter to the parents of shadow youth at 6 months after the Wave 2 parent contact and telephone the parents at approximately 12 months after that contact;
- Visit respondents who have moved up to 100 miles from a study PSU;[13] and
- When feasible, attempt to visit respondents who have moved more than 100 miles from a study PSU within the U.S.

[13] The proportion of SPs who move annually beyond 100 miles from any of the PSUs is estimated to be less than 1 percent.

Ongoing tracking of study respondents is essential to longitudinal cohort studies for purposes of cohort retention and follow up. Management of participant tracking and tracing activities by the PATH Study is through a centralized Home Management System (HMS). This component of the study management system houses the database of contact information, and it provides for real-time access in the field and at the home office to the most current information available. PATH Study staff involved in tracking and tracing activities provide updates to the HMS, and supervisors generate reports from the HMS to monitor progress in the field and identify the need for potential corrective actions.

The centralized HMS tool facilitates routine tracking steps that help to minimize the number of cases requiring intensive tracing. These steps include:

- **Collect contact information at the initial interview for tracing references.** At their initial PATH Study interview, respondents are asked for the names, addresses, and telephone numbers of two people who do not live in the same household and can serve as tracing references for how to reach the respondent. Given that a sizeable percentage of respondents are young adults who tend to be mobile, respondents are asked for additional information that may help to locate them (e.g., recent college attended).

- **Use interim contacts to determine if contact information has changed or if tracing is needed.** Contacts by mail ask respondents to report any contact information changes. The Study provides several ways this can be done, including visiting the Study website, calling a toll-free number, or sending updated information via mail. The PATH Study also mails materials to respondents stamped “return service requested,” requesting new address information for people who may have moved. In addition to supporting tracing, these interim contacts help to maintain respondent motivation to cooperate and continue engagement with the study. PATH Study respondents are offered an incentive ($5) as a thank you for updating their contact information on up to two occasions during the year, for a total of $10.

- **Update contact information as part of data collection.** During initial telephone contacts with households and household visits for each follow-up wave, the field interviewers update contact

information for adults as well as for relatives or persons not living in the household who can serve as references on where to locate the respondent.

The PATH Study's approach for tracing and locating respondents who may be lost at the Wave 3 follow-up is as follows: If current occupants of the respondent's last known address are unable to guide the field interviewer to a respondent's whereabouts, the field interviewer implements the first line of tracing using readily available information, including the respondent's last known telephone number(s), tracing references, directory assistance, and neighbors to try to locate the respondent. If unsuccessful, the case is sent to the PATH Study home office for the second line of tracing using the following protocols:

- **Lexis Nexis.** This database, compiled from public records, can return respondent address histories and telephone numbers. Submissions are made at least quarterly, and the tracers review and follow up on the results.
- **Internet searches.** These searches include free and paid services. Examples of the services include online telephone directories and limited public information records.
- **In-person tracing.** As the need arises and as resources permit, in-person tracing (i.e., "skip tracing") may be used. This approach involves intensive in-person tracing at the respondent's last known addresses and in his/her old neighborhoods to identify contact information or current location; in-person tracing differs from the first line of tracing by using specialists who develop leads that extend beyond those based on readily available information. Given its expense on a per case basis, in-person tracing is used rarely, after exhausting other approaches.

B.3b Wave 3 Data and Biospecimen Collection

To minimize attrition and maximize response rates in advance of Wave 3 data and biospecimen collection (and of each potential follow-up wave thereafter), the PATH Study employs field staff (interviewers and supervisors) who live within or in close proximity to the PSUs of SPs' homes. This helps to ensure that field staff are familiar with the communities within which their

assigned cases are located. Field interviewers are also trained in effective techniques for developing rapport and gaining respondent cooperation through refusal aversion and conversion.

In addition to the respondent incentives described in Section A.9, the PATH Study uses several tools and approaches to address nonresponse and maximize response rates. These include the following:

- The interviews are conducted in English and Spanish; all of the instruments are translated into Spanish, and bilingual field staff administer them.
- Materials for study respondents are designed to be informative and to encourage participation; all of the materials are translated into Spanish. These include follow-up or reminder letters that are sent 3 months prior to the household anniversary month to inform adult SPs and the parents of youth SPs about the planned Wave 3 data collection (Attachment 9). The letters remind recipients about the PATH Study's objectives, how its data will be used, why the study is important, and why the study includes tobacco users and non-users.
- Respondents can easily access information about the PATH Study through the PATH Study website and a toll-free respondent telephone call line dedicated to answering respondents' questions and verifying the credibility of the study.

Approximately 1 month prior to the household anniversary month, field interviewers telephone the adult SPs and parents of youth SPs. This call helps to reestablish direct contact, collect updated contact information, identify the parent who will participate for each youth SP or shadow youth, answer questions about the Study, and make an appointment for the in-person visit. As needed (e.g., several telephone contact attempts are unsuccessful), the field interviewers make the first direct contact in-person.

Refusal letters and follow-up contacts are used with adults who are reluctant to participate and/or are located in limited-access situations (e.g., doorperson buildings and gated communities). The letters are tailored to the reason for reluctance or refusal (e.g., too busy or concerned about confidentiality) and to the type of respondent (e.g., adult SP or parent of

youth SP). (See Attachment 19 for an example of a refusal letter.); these letters may be sent via FedEx or priority mail to reinforce the perceived importance of participation. Also, email messages based on refusal letters may be sent to refusing or difficult-to-contact adults who agreed to be contacted this way. After the letters are sent to respondents, field interviewers work with their supervisors to tailor the approach for making follow-up contacts with the respondents by telephone or in person. Following best practice, the field interviewer making the follow-up contact is often different from the field interviewer who made the original contact.

Additional tools and approaches will be used by the PATH Study to help maximize the biospecimen response rates for Wave 3, including the following.

- **Ensure that Interviewers are "On Board."** The PATH Study continues to hire and train interviewers who understand the importance of collecting biospecimens as part of this research effort. Early in the selection process, candidates are required to view a short video that highlights this requirement and the importance of being comfortable with carrying it out.

- **Phase the Consent for Biospecimens.** For youth SPs who age up to the adult cohort, the PATH Study presents information to respondents in phases to help minimize the amount of information to be simultaneously considered before consenting. This approach includes providing information about the interview immediately prior to obtaining consent for the interview; providing information on biospecimen collection shortly before obtaining consent for biospecimen collection, etc. Moreover, because biospecimen collection follows completion of the interview, this approach allows additional time for the development of rapport, trust, and comfort between the interviewer and the respondent, which positively influence consent to provide the biospecimens.

- **Present the Biospecimens in a Positive Light.** Based on an effective approach used by the National Health and Nutrition Examination Surveys (NHANES), the PATH Study uses nicely-formatted consent pamphlets with messages that emphasize the importance of the respondent's contributions of biospecimens to the Study's scientific success.

- **Enhance Training of Interviewers.** The PATH Study continues to provide extensive interviewer training on collecting biospecimens, including home-study training and practice in requesting consent and averting refusals. With classroom and home-study training and additional practice sessions, interviewers are able to gain proficiency and comfort with the study protocol, including obtaining consent, averting refusals, and collecting biospecimens.

- **Equip Interviewers with Refusal Conversion Tool.** The PATH Study continues to use computer-assisted personal interviewing (CAPI) screens that, in real time, point interviewers to tailored responses to types of reasons respondents give for biospecimen refusals. Having these available at the moment they are needed can improve the interviewer's ability to quickly allay respondent concerns about providing biospecimens.

- **Streamline Biospecimen Collection Procedures.** The PATH Study follows the same procedure that was adopted for Wave 1 and used in Wave 2 of having the field interviewer collect a urine sample at the time of the interview for participants providing urine at Wave 3. Rapport that develops between the interviewer and respondent appears to have a positive influence on the respondent's willingness to provide the biospecimens. As noted, these participants will be the youth who age up to the adult cohort at Wave 3 and a subsample of approximately 17,386 adults who initially provided urine at Wave 1 or Wave 2 and are asked to provide urine at Wave 3.

- **Enhance Quality Control.** The PATH Study's data collection quality control procedures include closely monitoring interviewer-by-interviewer consent and collection rates for biospecimens, using computer-assisted recorded interviewing (CARI) to monitor interviewer performance on the consent and collection tasks, and providing rapid feedback to interviewers and refresher training to maximize performance.

A web-based Supervisor Management System (SMS) allows field supervisors to monitor each field interviewer's work and help in the development of strategies to address nonresponse. These strategies may include reassigning difficult or reluctant cases among local field interviewers; and using specially-trained, traveling field interviewers with experience in refusal conversion.

Data collection efforts also follow a phased approach that anticipates refusal conversion efforts. In this approach, samples of SPs are released to field interviewers in sets every few months; the timing of the releases is tied to the household anniversary of the SPs. Closing out cases from an earlier set is not necessary before releasing cases in a new set, thus allowing additional time to complete challenging cases. Further, the number of cases assigned to interviewers is expected to be lowest during later periods in the data wave, thereby ensuring interviewers have additional time in those periods to complete open cases remaining from an earlier period. “Front loading” the sample releases in this manner allows field interviewers the opportunity to implement the full contact strategy, including nonresponse conversion as needed.

Adjustments will be performed as necessary for non-interviews that cannot be converted using the procedures described in Section B.2. The specific procedure selected ensures the accuracy of resulting estimators and the suitability of the compensated data set for addressing the major objectives of the study.

B.4 Test of Procedures or Methods to Be Undertaken

The PATH Study Wave 2 data and biospecimen collection, which is currently underway, serves as an informal test of many of the methods and materials planned for Wave 3. This is reasonable given that many of the methods and materials in Wave 2 will be similar for Wave 3. For example, for the cohort movers (youth SPs who age up to the adult cohort and shadow youth who age up to the youth cohort), the Wave 3 consent, parental permission, and youth assent procedures and methods are like those used in Wave 2. The Wave 3 methods for administering the ACASI and CAPI instruments, collecting biospecimens, and paying incentives for all types of participants are also similar to the methods used for Wave 2.

In addition to this informal test, the PATH Study developed the afore-mentioned Interim Report based on the complete Wave 1 data and biospecimen collection and on approximately the first 6 months of Wave 2 data and biospecimen collection. Separately for these two waves, the report includes the actual or projected response rates (for screening at Wave 1, interview completion, and biospecimen collection); the results of nonresponse analysis and the plan for future statistical analyses; and the implications of the response rates and nonresponse bias for the types of conclusions that can be drawn from this study. Response rates are compared throughout this report to corresponding rates projected for the entire sample, provided in Supporting Statement B of the PATH Study's non-substantive change requests for Wave 1 of data and biospecimen collection and for Wave 2 of data and biospecimen collection.

The PATH Study's Wave 2 Interim Report is provided in Attachment 21. Select findings on Wave 1 and Wave 2 are summarized in the remainder of this section.

B.4a Wave 1 Data and Biospecimen Collection

Response Rates for Wave 1

Response rates, including response rates weighted with inverse probabilities of selection (IPS), were computed for the three data collections and two biospecimen collections.[14] The response rates for the collections vary on how they compare to the rates projected in the Wave 1 non-substantive change request to OMB and "worst-case" scenario rates for the full sample provided in Attachment 22. (Please see the introduction and reference to Attachment 22 in the Interim Report for more information on these scenarios.)

[14] Buccal cells were collected from adult participants for approximately 8 of the 15 months of Wave 1 (through May 18, 2014, when OMB approved a change request to discontinue the collection of this biospecimen). The PATH Study did not resume buccal cell collection in Wave 2; in addition, the Study has no plans to resume this collection in Wave 3.

- The weighted response rates for the Household Screener (54.0%) and Adult Extended Interview (74.0%) are lower than the projected rates for the full sample, but they exceed the worst-case scenario rates. The weighted predicted response rate for the Youth Interview (78.4%) exceeds the projected rate.

- The weighted response rates for the biospecimen collections—urine (63.6%) and blood (43.0%)—are lower than the projected rates. These rates exceed the worst-case scenario rates.

- As discussed further in the report, differences among tobacco use status and demographic subgroups on response rates for the collections are generally modest.

Nonresponse Bias Analysis for Wave 1

PATH Study IPS-weighted estimates were compared with estimates from national cross-sectional studies. Most of the PATH Study estimates are consistent with those from other studies.

- The PATH Study estimates of percentage of single- and two-person households are lower than those in the 2013 American Community Survey (ACS). The estimated percentage of persons who are non-Black and 25 years of age or older, from the household rosters, is also smaller than the corresponding estimate from the ACS.

- Hispanics, persons in the "other race" category, adults 18 to 24 years old and 25 to 44 years old, and males are somewhat over- or under-represented in the PATH Study estimates for some specific data or biospecimen collections.

- The PATH Study's estimates of adult cigarette smoking are in line with those from other studies. Its estimates of youth cigarette smoking are toward the low end of the range of estimates found for other studies.[15]

[15] Estimates of adult cigarette smoking from the PATH Study were compared with estimates from the Tobacco Use Supplement to the Current Population Survey, 2010-2011 (TUS-CPS); the National Health and Nutrition Examination Survey, 2011-2012 (NHANES); the National Health Interview Survey, 2013 (NHIS); and the National Survey on Drug Use and Health, 2013 (NSDUH). Estimates of youth cigarette smoking from the PATH Study were compared with estimates from NHANES, NSDUH, and the National Youth Tobacco Survey, 2012 (NYTS). Results from the 2013 NSDUH (SAMHSA, 2014) indicate that youth cigarette smoking dropped from between 0.8 to

Approach to Address Nonresponse for Wave 1

The statistical approach to address nonresponse is to adjust the IPS weights to account for nonrespondents. This approach was successful overall in correcting for nonresponse bias on characteristics measured in the ACS.

- Applying the adjusted IPS weights to the full sample reduces the discrepancy between the PATH Study estimates and the 2013 ACS estimates on demographic characteristics.
- Estimates of adult cigarette smoking using the adjusted weights remain in line with those from other surveys. Estimates of youth smoking using the adjusted weights remain at the low end compared with other surveys.

B.4b Wave 2 Data and Biospecimen Collection

Predicted Response Rates for Wave 2

The PATH Study Wave 2 data collection is ongoing, so response rates were calculated on a probability subsample of cases using predicted response propensities for nonfinalized or interim cases. The predicted rates are inconclusive, because only 76 percent of the cases examined have been finalized to date and the response status is predicted for the remaining 24 percent using statistical models. The predicted response rates were computed for continuing adults and continuing youth as well as for aged-up adults and aged-up youth (those who, in Wave 2, became age-eligible and participated in the PATH Study for the first time). The response rates for the collections vary on how they compare to the rates projected in the Wave 2 Revision Request to OMB.

- The predicted weighted response rates for interviews for continuing adults (84.3%), continuing youth (89.2%), and aged-up adults (86.4%) are slightly lower than the projected rates. The weighted predicted response rate for the aged-up youth (87.5%) exceeds the projected rate.

2.7 percentage points for various age subgroups between 2012 and 2013.

- For continuing adults, the weighted predicted response rate for urine collection (96.1%) exceeds the projected rate. For aged-up adults, the weighted predicted rate for urine (82.3%) exceeds the projected rate, but the predicted rate for blood (43.2%) is slightly lower than the projected rate.

- The predicted response rates exhibit little variability across population subgroups.

Nonresponse Bias Analysis for Wave 2

- The nonresponse bias analysis suggests that estimated percentages of persons with less than a high school degree tend to be lower for the Wave 2 respondents than for the Wave 2 finalized nonrespondents and provisional nonrespondents (finalized nonrespondents plus interim refusals and persons who are difficult to locate), although the difference between the respondents and finalized nonrespondents is not statistically significant.

- Wave 1 tobacco use rates are not significantly different between Wave 2 respondents and finalized nonrespondents, although the provisional nonrespondents exhibit higher tobacco use rates than the respondents.

- Wave 2 provisional nonrespondents also have lower health insurance coverage (at Wave 1) than respondents, although the difference between respondents and finalized nonrespondents is not significant.

Approach to Address Nonresponse for Wave 2

For Wave 2, weights of respondents will be adjusted to account for nonrespondents by forming weighting adjustment cells using Wave 1 characteristics of respondents and nonrespondents. Consequently, nonresponse-adjusted weights of Wave 2 respondents will sum to Wave 1 population totals. This weighting will compensate for differences between respondents and nonrespondents with respect to sex, age, other demographic variables, and Wave 1 tobacco use status.

B.5 Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

A list of individuals who consulted on statistical aspects of the PATH Study design and will collect and/or analyze the PATH Study data is included in Attachment 22.

References

Brick, J.M. (2013). Unit nonresponse and weighting adjustments: A critical review. *Journal of Official Statistics*, *29*, 329–374 (with discussion).

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics,* 19, 473–503.

De Waal, T., Pannekoek, J., & Scholtus, S. (2013). *Handbook of Statistical Data Editing and Imputation.* Hoboken, NJ: Wiley.

Guilliford, M.C., Ukoumunne, O.C., & Chinn, S. (1999). Components of variance and intraclass correlations for the design of community-based surveys and intervention studies. *American Journal of Epidemiology*, 149(9), 876–883.

Judkins, D., Krenzke, T., Piesse, A., Fan, Z., & Haung, W.C. (2007). Preservation of skip patterns and covariate structure through semi-parametric whole questionnaire imputation. *American Statistical Association Proceedings of the Survey Research Methods Section*: 3211–3218.

Kalton, G. (1986). Handling wave nonresponse in panel surveys. *Journal of Official Statistics, 2*, 303–314.

Kashihara, D. & Ezzati-Rice, T. (2004). Characteristics of survey attrition in the household component of the Medical Expenditure Panel Survey. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 3758–3765.

Lipsey, M. (1990). *Design Sensitivity: Statistical Power for Experimental Research*. Newbury Park, CA: Sage.

Lynn, P. (2006). Quality Profile: British Household Panel Study. Essex, UK: Institute for Social and Economic Research. Retrieved from https://www.iser.essex.ac.uk/files/bhps/quality-profiles/BHPS-QP-01-03-06-v2.pdf

Micklewright, J., Schnepf, S., & Skinner, C. (2012). Non-response biases in surveys of schoolchildren: The case of the English Programme for International Student Assessment (PISA) samples. *Journal of the Royal Statistical Society, Series A*, 175, 915–938.

National Institutes of Health (2010). Alcohol use and alcohol use disorders in the United States, a 3-year follow-up. U.S. Alcohol Epidemiologic Data Reference Manual, Volume 8, Number 2. Retrieved from pubs.niaaa.nih.gov/publications/NESARC_DRM2/NESARC2DRM.pdf

National Longitudinal Survey of Youth, 1997 (2014). Retention & Reasons for Non-Interview. Retrieved from https://www.nlsinfo.org/content/cohorts/nlsy97/intro-to-the-sample/retention-reasons-non-interview/page/0/0/#retention

National Research Council (2014). *Nonresponse in Social Science Surveys: A Research Agenda.* Washington, D.C.: National Academies Press.

Roth, S., Montaquila, J., & 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.

Schouten, B. & de Nooij, G. (2005). *Nonresponse adjustment using classification trees.* Discussion paper 05001, Statistics Netherlands. Available at www.cbs.nl.

Substance Abuse and Mental Health Services Administration. (2014). Results from the 2013 National Survey on Drug Use and Health: Summary of national results. Retrieved from http://www.samhsa.gov/data/NSDUH/2013SummNatFindDetTables/NationalFindings/NSDUHresults2013.htm

Särndal, C.E. & Swensson, B. (1987). A general view of estimation for two phases of selection with applications to two phase sampling and non-response. *International Statistical Review*, 55, 279-294.

Statistics New Zealand (2011). New weighting methodology for longitudinal surveys: As applied in the Survey of Family, Income and Employment. Retrieved from http://www.stats.govt.nz/browse_for_stats/income-and-work/Income/sofie/new-weighting-methodology-for-sofie/response-rates.aspx

Thompson, D.M., Fernald, D.H., & Mold, J.W. (2012). Intraclass correlation coefficients typical of cluster-randomized studies. *Annals of Family Medicine*, *10,* 235–240.

Watson, N. & Wooden, M. (2006). Modelling longitudinal survey response: The experience of the HILDA Survey. Paper presented at the ACSPRI Social Science Methodology Conference, Sydney Australia. Retrieved from

http://conference2006.acspri.org.au/proceedings/streams/Modelling%20Longitudinal%20Survey%20Response.pdf

Watson, N. & Wooden, M. (2011). Re-engaging with survey nonrespondents: The BHPS, SOEP, and HILDA experience. Melbourne Institute Working Paper Series no. 2/11. Melbourne: Institute of Applied Economic and Social Research, University of Melbourne. Retrieved from https://www.melbourneinstitute.com/downloads/hilda/Bibliography/HILDA_Discussion_Papers/hdps111.pdf