Attachment 22

Precision Calculations under Projected (Best-Case) and Worst-case Scenarios

June 18, 2013

Precision Calculations under Projected (Best-Case) and Worst-Case Scenarios

Table 1 presents the expected sample sizes under three response-rate scenarios. The first set of response rates, in columns 2 and 3, are from the original OMB submission; these are reproduced here for purposes of comparison with the original document. The second set of response rates, in columns 4 and 5, are the projected "best-case" response rates of 70 percent to the household screener, 85 percent for adults completing the second-phase sampling questions, and 65 percent for blood collection; these are the projected rates presented in Table 2 of Supporting Statement B (SSB)[1]. The third set of response rates, in columns 6 and 7, presents the unweighted response rates observed from the field test, and represents the "worst-case" scenario at every step of data collection: a 39.7 percent response rate to the household screener, a 58.1 percent response rate for adults completing the second-phase sampling questions, and a 39 percent response rate for collection of blood samples.

Despite the 39 percent response rate for blood samples achieved in the field test, there are reasons to have confidence in the projected 65 percent response rate for blood collection for the main study, as presented in SSB. First, compared to the field test, the main study will have a longer data-collection period which will afford more opportunities to follow-up with respondents; this is projected to reduce the number of persons consenting but not providing blood samples (which occurred in many cases in the field test because the data-collection period ended abruptly, before blood samples could be collected from many respondents who consented to provide them). Second, based on the field test, the PATH Study identified a number of improvements to make in its blood-collection procedures (see Section B.3 of SSB). Third, past and recent experience in the collection of blood specimens in with blood collection in other studies (e.g., the Kidney Center Study, which achieved conditional blood response rates of 68 percent for African-Americans and 73 percent for whites; and the National Health and Nutrition Examination Surveys, which achieved weighted conditional blood response rates of 83 percent or higher for all adult age groups) suggests that a 65 percent blood-collection response rate is both reasonable and achievable in the PATH Study.

Table 2 presents the expected sample sizes for the number of blood samples obtained overall and for various subgroups under the projected response rate and the worst-case scenario. Blood collection is the sample component with the lowest expected response rate and sample size;

[1] Note that the number of addresses sampled has been increased to account for the anticipated 70 percent response rate to the household screener; as a consequence, the anticipated number of adults sampled is the same for the projected response rates as in the original OMB submission.

therefore, expected precisions for all other aspects of data collection—responses to the adult questionnaire, buccal cell, and urine—will be higher than that for the blood collection. Precision estimates for blood collection, then, are worst-case precisions for all the items in the survey. All calculations assume the within-household adult subsampling rates used to obtain the relative weights in Table 1b of SSB, as well as a similar cooperation rate for blood collection across demographic/user groups. Table 2 also presents relative standard errors for a dichotomized variable assumed to have 15 percent prevalence. For the anticipated response rates of 70 percent to the household screener and 65 percent for blood collection, the relative standard errors for this dichotomized variable are below 10 percent for all but the smallest population subgroups, and are below 5 percent for menthol smokers, 18-24-year-old current users, daily users, and black/African American adults. The relative standard errors are approximately twice as large under the worst-case scenario. Yet even under the worst-case scenario, the relative standard error is below 5 percent for the larger subgroups of adult tobacco users and adults age 25 and over, and the relative standard error is below 10 percent for subgroups of interest such as menthol users, experimental tobacco users, and 18-24-year-old current tobacco users.

Table 3 presents the minimum standardized detectable differences (MSDDs) for selected comparisons on a continuous variable under the projected response rate and the worst-case scenario, assuming that 80 percent power is desired. Comparisons presented in Table 3 center around differences between daily/less than daily/non-tobacco users and menthol/non-menthol cigarette smokers for all adults and for young adults. The third column of Table 3 presents the MSDDs assuming the anticipated 70 percent household screener response rate: the MSDDs for the selected comparisons are all less than 0.07 for comparisons among all adults, and less than 0.12 for comparisons within the smaller 18-24 year-old group. For the worst-case scenario of a 39.7 percent household screener response rate and a 39 percent response rate for blood collection, the MSDDs presented in the fourth column are less than 0.15 for all adults, and 0.25 for 18-24 year-old adults. Thus even under the worst-case scenario, the MSDDs are in the range generally considered to be small effects from an empirical perspective.[2]

In summary, the relative standard errors and MSDDs associated with the anticipated sample sizes from the PATH Study are sufficiently small to allow detection of differences in biomarkers from blood samples that are of importance to assess tobacco exposure and indicators of disease

[2] Lipsey, M. (1990). Design Sensitivity: Statistical Power for Experimental Research. Newbury Park, CA: Sage. Lipsey considers an effect of 0.15 standard deviations to be "small," and an effect of 0.45 standard deviations to be "medium."

risk. Even under the "worst case" scenario of a 39.7 percent response rate to the household screener and a 39 percent response rate to the blood collection, the PATH Study will have sufficient precision on dichotomized variables for many population subgroups of interest, and will be able to detect small differences between population subgroups on continuous variables.

Table 1. **Response rate assumptions for the PATH Study at baseline. The first two columns are from Table 2 of Supporting Statement B, and assume the projected ("best-case scenario") 70% household screener response rate. The last two columns are computed assuming a 39.7% household screener response rate and a 58.1% extended interview response rate.**

Sampling Unit	Assumed rate, Original SSB	Expected number, Original SSB	Assumed rate, Revised SSB with projected 70% household screener response rate	Expected number, Revised SSB with projected 70% household screener response rate	Assumed rate, Revised SSB with worst-case scenario of 39.7% household screener response rate	Expected number, Revised SSB with worst-case scenario 39.7% household screener response rate
Primary sampling unit (PSU)	——	150	——	156	——	156
Area segments/CDSF segments	40 per PSU	6,000	40 per PSU	6,000	40 per PSU	6,000
Addresses	22.1 per segment	132,668	28.1 per segment	168,857	28.1 per segment	168,857
Occupied dwelling units	88.6%	117,544	88.60%	149,607	88.60%	149,607
Households completing screener enumeration	87%	102,263	70%	104,725	39.7%	59,394
Eligible households with adults	100%	102,263	100%	104,725	100%	59,394
Number of adults sampled at first stage	Up to 2 per HH	70,000	Up to 2 per HH	70,000	Up to 2 per HH	39,700
Number of adults completing second-phase sampling questions at beginning of extended interview	90%	63,000	85%	59,500	58.1%	23,066
Number of adults retained at second phase of sampling and completing full extended interview	68%	42,730	72%	42,730	72%	16,565
Number of adults completing extended interview who provide buccal cells	85%	36,321	80%	34,184	73%	12,092

Number of adults completing extended interview who provide urine	85%	36,321	80%	34,184	49%	8,075
Number of adults completing extended interview who provide blood	65%	27,775	65%	27,775	39%	6,460
Number of adults completing extended interview who provide all biospecimens	65%	27,775	65%	27,775	39%	6,460

Table 2. **Relative standard errors at baseline from blood collection for item with 15 percent prevalence. The first two columns assume a projected ("best-case scenario") 70% response rate for the household screener and a 65% response rate for blood collection; the last two columns assume a 39.7% response rate for the household screener and a 39% response rate for blood collection.**

Group	Baseline sample size, assuming 70% household screener response rate and 65% response rate for blood collection	RSE (%) on 15% item, assuming 70% household screener response rate and 65% response rate for blood collection	Baseline sample size, for worst-case scenario with 39.7% screener response rate and 39% response rate for blood collection	RSE (%) on 15% item, for worst-case scenario with 39.7% screener response rate and 39% response rate for blood collection
All adults	**27,775**	**2.8%**	**6,460**	**4.4%**
Current users	12,930	2.9%	3,008	4.9%
Menthol smokers	3,750	4.5%	872	8.6%
Dual (smokers and smokeless tobacco users)	1,930	6.0%	449	11.8%
Experimental tobacco users	2,995	4.9%	697	9.5%
Daily users	10,344	3.1%	2,406	5.4%
Less-than-daily users	2,586	5.2%	601	10.2%
Current non-users	7,140	3.6%	1,661	6.4%
Adults ages 18-24	**6,961**	**3.8%**	**1,619**	**6.9%**
Current users	3,080	4.7%	717	9.1%
Menthol smokers	1,048	7.6%	244	15.3%
Dual (smokers and smokeless tobacco users)	757	8.8%	176	18.0%
Daily users	2,218	5.4%	516	10.6%

Less-than-daily users	863	8.3%	201	16.8%
Current non-users	1,338	6.8%	311	13.7%
Adults ages 25+	**20,814**	**2.9%**	**4,841**	**4.8%**
Current users	9,850	3.2%	2,291	5.4%
Current non-users	5,801	3.7%	1,349	6.8%
Black/African American adults	**3,900**	**4.9%**	**907**	**9.6%**
Current users	1,806	6.1%	420	12.1%
Menthol smokers	1,319	7.0%	307	14.1%
Dual (smokers and smokeless tobacco users)	270	15.1%	63	31.0%
Daily users	1,336	7.0%	311	14.1%
Less-than-daily users	470	11.5%	109	23.6%
Current non-users	1,180	7.5%	275	15.0%

Table 3. **Minimum standardized detectable differences (MSDDs) with 80 percent power, for selected comparisons. The first column assumes a projected (or best-case scenario)70% response rate for the household screener and a 65% response rate for blood collection; the last column assumes a 39.7% response rate for the household screener and a 39% response rate for blood collection.**

Group 1	Group 2	MSDD, 65% response rate for blood collection, assuming 70% HH screener response rate	MSDD, 39% response rate for blood collection, assuming 39.7% HH screener response rate
Adult daily users	Adult less-than-daily users	0.062	0.206
Adult daily users	Adult current non-users	0.043	0.089
Adult menthol smokers	Adult current non-users	0.057	0.117
Adult menthol smokers	Adult current users, non-menthol smokers	0.054	0.113
Age 18-24 daily users	Age 18-24 current non-users	0.097	0.201
Age 18-24 menthol smokers	Age 18-24 current non-users	0.116	0.240
Age 18-24 menthol smokers	Age 18-24 current users, non-menthol smokers	0.107	0.221