

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

Section B

B.1. Respondent Universe and Sampling Methods

There are three populations of interest under the Peer Review Enhancement Surveys: an applicant population, reviewer population and Advisory Council member populations. These populations are defined as follows:

Applicant Population

The applicant population comprises those individuals who submitted R01, R03, R21, U01 and R34 applications to NIH reviewed in any of the Advisory Councils/Boards of NIH's constituent Institutes and Centers (ICs) in May 2011 or October 2011.

Reviewer Population

The reviewer population comprises those individuals who served in NIH study sections that reviewed R01, R03, R21, U01 and R34 applications that were subsequently reviewed by the Advisory Councils/Boards in October 2011, and/or February 2012. The target population of reviewers includes regular (appointed/permanent) and *ad hoc* (temporary) reviewers.

Advisory Council Member Population

The Advisory Council member population includes members who are regular (chartered) members and served at one of the NIH National Advisory Council/Board meeting held in October 2011 or February 2012. All Advisory Council members will be invited to take the Advisory Council survey. We anticipate that the number of members who will be eligible to take the Advisory Council survey will be 250.

Applicant and Reviewer Population

There are some individuals who are eligible to be members of both the applicant population and in the reviewer population. The sampling design for the peer review surveys was developed so that no individual who resides in both populations would be contacted for both the Applicant Survey and the Reviewer Survey. Table B.1-1 shows the total number of individuals in the universe of all applicants and reviewers (column 2), the number of individuals who are applicants but not reviewers (column 3) and the number of individuals who are reviewers but not applicants (column 4). Table B.1-1 also shows the numbers of individuals by race and ethnicity¹ in the total applicant population (column 7), in the total reviewer population (column 6), and in the total population of individuals who are both an applicant and a reviewer (column 5).

¹ 1,746 Individuals with Unknown Hispanicity were assumed to be non-Hispanic for purposes of sample selection. 7,523 Individuals with Unknown Race were included in the "Other" race category along with 29,504 Whites and four multi-racial individuals who were not classifiable as Asian, Black, Native American, or Pacific Islander. Four Hispanic, Pacific Islanders were classified as Hispanic, Other for purposes of sample selection because of the extremely limited number of individuals in this group.

It also is possible for an Advisory Council member to be an Applicant and in some rare cases a Reviewer. Any eligible Advisory Council members who appear in the Applicant and Reviewer sampling frame will be removed from the sampling frame prior to drawing the sample of Applicants and Reviewers who will be invited to participate in the Applicant Survey and the Reviewer Survey.

The total number of applicants and reviewers (37,843) is equal to the sum of the number of individuals who are applicants only (24,385), the number of individuals who are reviewers only (7,980), and the number of individuals who are both applicant and reviewer (5,478). The number of individuals who are applicants (29,863) equals the number of individuals who are applicants only (24,385) plus the number of individuals who are both applicant and reviewer (5,478). The number of individuals who are reviewers (13,458) equals the number of individuals who are reviewers only (7,980) plus the number of individuals who are both an applicant and a reviewer (5,478).

Table B.1-1. Applicant and Reviewer Population Counts

	(2)	(3)	(4)	(5)	(6)	(7)
Stratum	Total of Applicants and Reviewers	Applicants Only	Reviewers Only	Both Applicant and Reviewer	Total Applicant Population	Total Reviewer Population
American Indian/ Alaska Native, Hispanic	23	20	2	1	21	3
Asian, Hispanic	19	14	1	4	18	5
Black, Hispanic	16	10	5	1	11	6
Multiracial, Hispanic	45	25	11	9	34	20
Other, Hispanic	1,288	779	305	204	983	509
American Indian/ Alaska Native, non-Hispanic	43	25	11	7	32	18
Asian, non-Hispanic	6,408	4,496	924	988	5,484	1,912
Black, non-Hispanic	655	410	162	83	493	245
Multiracial, non-Hispanic	272	171	55	46	217	101
Other, non-Hispanic	29,051	18,423	6,497	4,131	22,554	10,628
Pacific Islander, non-Hispanic	23	12	7	4	16	11
Total	37,843	24,385	7,980	5,478	29,863	13,458

Note: A total of 13,206 persons with unknown ethnicity were assumed to be non-Hispanic for purposes of sample selection. 6,406 persons with unknown race were included in the “other” race category, together with 23,933 Whites.

Sample Selection

Determining Overall Sample Sizes

The total number of individuals who may be sampled and subsequently surveyed is defined by burden limits under NIH’s Office of Management and Budget (OMB) Generic Clearance No. 0925-0627. For the applicant and reviewer surveys, the total number of persons who may be sampled under the burden limits established by the NIH guidance is 4,460. Given the total number of allowable sample members, the next step is to determine how many of the allowable 4,460 should be selected from the applicant population and how many should be selected from the reviewer population.

Broad Allocation Scheme

The total number of individuals to be sampled will be allocated to the following sets of individuals:

1. Applicants only (24,385)
2. Reviewers only (7,980)
3. Those who are both Applicant and Reviewer (5,478)

Within each set, sample sizes must be sufficient in order to allow for estimates within race and ethnicity groups to meet precision requirements (discussed below).

Initial Sample Sizes Based on Precision Requirements

The following four steps were taken for the three groups of individuals: Applicants only, Reviewers only, and those who are both Applicant and Reviewer.

The following four steps will be taken for the three groups of sample members—applicants only, reviewers only, and those who are both reviewers and applicants:

1. A cross-tabulation will be created of the number of individuals by race (Asian, Black, Native American, Pacific Islander, “other,” and multiracial) and ethnicity (Hispanic or non-Hispanic).
2. Using the nQuery Advisor software (Elashoff, 2005), the number of individuals required to be sampled in each race-by-ethnicity group will be estimated such that, within each group, a two-sided 95% confidence interval for a population proportion of 50% will have a half-width of 5%.
3. For those groups with population counts of less than 30 or for which nQuery reports that the sample size is not large enough for the sample calculation to be approximately normally distributed, all group members in the relevant sample will be included. Such groups are said to be selected with certainty.
4. For those groups not selected with certainty, nQuery will report the required sample size.

The following table (Table B.1-2) shows the number of individuals selected with certainty or estimated by nQuery as being required to meet the precision requirement outlined in Step 2 above.

Table B.1-2. Initial Sample Sizes Based on Precision Requirements

Stratum	Applicants Only		Applicants and Reviewers		Reviewers Only	
	(1) Population Count	(2) Sample Size	(3) Population Count	(4) Sample Size	(5) Population Count	(6) Sample Size
American Indian/ Alaska Native, Hispanic	20	20	1	1	2	2
Asian, Hispanic	14	14	4	4	1	1
Black, Hispanic	10	10	1	1	5	5
Multiracial, Hispanic	25	25	9	9	11	11
Other, Hispanic	779	258	204	204	305	171
American Indian/ Alaska Native, non-Hispanic	25	25	7	7	11	11
Asian, non-Hispanic	4,496	354	988	554	924	272
Black, non-Hispanic	410	199	83	83	162	114
Multiracial, non-Hispanic	171	119	46	46	55	55
Other, non-Hispanic	18,423	377	4,131	705 ¹	6,497	363
Pacific Islander, non-Hispanic	12	12	4	4	7	7
Total	24,385	1,413	5,478	1,618	7,980	1,012

The sample sizes listed in Column 4 will be subsampled such that half of the individuals selected will be assigned the applicant questionnaire; half, the reviewer questionnaire. The sample sizes listed in Column 4 for those race-ethnicity groups not selected with certainty are designed so that the precision requirements will still be met for the 809 (half of 1,618) individuals selected to receive the applicant questionnaire and for the 809 others selected to receive the reviewer questionnaire.

Column 5 shows the number of individuals who are reviewers only, and Column 6 lists the numbers of individuals who must be sampled to achieve the stated precision requirement.

After determining the sample sizes required to meet the stated precision requirement, the total number of individuals required to be sampled is $1,413 + 1,618 + 1,012 = 4,043$. Because the burden requirement allows for a total of 4,460 individuals to be sampled, $4,460 - 4,043 = 417$ more individuals may be sampled.

Allocating the Remaining Sample of 417 by Consideration of Weighting

The remaining 417 individuals will be allocated to the three possible samples according to the impact of sampling weights on the precision of estimates generated from each of the samples. A widely used measure for assessing the degree to which sampling weights affect the precision of statistical estimates is known as the *design effect*. The design effect for a given statistical estimate is the ratio of the variance of the estimate under the appropriate complex sampling process to the variance of the estimate when it is assumed that the underlying data arose by means of a simple random sample. Because as many design effects exist as potential estimates, a particular approximation of a design effect is used in practice. It is defined as follows: Given a sample of n individuals with a set of associated sampling weights denoted w_i and with the average sampling weight denoted W , then

$$\text{Deff} = 1 + \left(\frac{1}{n} \right) \left(\frac{\text{Sum}(w_i - W)^2}{(W)^2} \right)$$

In the case of simple random sampling, the design effect is 1. When design effects are larger than 1, then the variances of estimates are larger than they would be if the samples had been selected with simple random sampling. A general guideline is to try to keep the Deff from exceeding 2. In accordance with this consideration, the additional 417 individuals will be allocated to reduce this design effect.

After allocating the additional 417 individuals, the design effects will be 1.62 for the applicant sample, 1.62 for the reviewer sample, and 1.42 for those individuals who were both applicants and reviewers. The final samples sizes based on precision requirements and design effect considerations are the same as shown in table B.1 -2 except for the Other, non-Hispanic stratum, which is 673 applicant-only, 705 applicant-reviewer and 484 reviewer-only.

Sample Power Analysis

Power estimates were calculated for comparing various race and ethnicity groups for each of the following samples: 1) Applicant only; 2) Reviewer only; 3) Applicant only plus those who are both an Applicant and a Reviewer and were selected to receive the applicant questionnaire; and 4) Reviewer only plus those who are both an applicant and a reviewer and were selected to receive the reviewer questionnaire. The power estimates ranged from a minimum of 10% to a maximum of 48% for detecting a difference of 5%. The power estimates ranged from a minimum of 533% to a maximum of 97% for detecting a difference of 10%.

Response Rates

The response rates for the survey will be calculated in accordance with the recommendations that the American Association for Public Opinion Research (AAPOR) has published in its *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (2008). The formula for the response rate is as follows:

$$RR6 = \frac{(I + P)}{[(I + P) + (R + NC + O)]},$$

where I = complete interview, P = partial interview, R = refusal, NC = noncontact, and O = other nonresponse. Notably, this formula differs from the AAPOR formula RR4 in that, because all individuals in the NIH-provided sampling frame are assumed to be eligible for the study, no estimate of the number of eligible individuals among those with unknown eligibility is included in the denominator. Adjustments to the response rate formula can be made if ineligibility of some individuals is later determined.

Sample Weights

Discussed here is the method to be followed to create the final sample weights and final estimates for the peer review surveys. One nonresponse-adjusted sample weight will be created for the applicant sample; another weight, for the reviewer sample. These weights will consist of a product of two factors: the base weight and the nonresponse adjustment, defined as follows:

- The *base weight* (for a given sample) is the inverse of the unconditional probability of selecting a sample member into the sample. This weight accounts for the stratification used in the sample design. Notably, if all sampled individuals respond, then no nonresponse adjustment is necessary.
- The *nonresponse adjustment* (for a given sample) is an adjustment imposed on the sampling weight of the respondents to account for those applicants who do not

respond to the survey. In general, this adjustment will be greater than 1 so that each respondent will represent himself or herself, as well as some portion of the nonrespondents.

There are numerous ways of constructing a nonresponse adjustment. For each of the applicant and reviewer samples, the plan is to adjust the base weights within strata and to use a simple ratio adjustment. In order to perform this adjustment, we will need to know which stratum each respondent belongs to.

Estimation Procedure

After the data are collected, analysis of the data must rely on software that can account for the sample design. Data analysis will be performed with SUDAAN software (2008). SUDAAN can manage correlated observations in a general sense, with nonparametric and parametric approaches being available. Base SAS software will be used for data manipulation and tabulation of results (2008).

B.2. Information Collection Procedures/Limitations of the Study

Data Collection Procedures

Sample members will be asked to complete the surveys online. The basic steps involved in the data collection process for all three surveys include:

- A lead letter will be sent to each sample member via email during the week prior to the day the survey opens (Attachment 6). The lead letter will be signed by a senior NIH official. It will explain the purpose of the survey and why they were selected to participate.
- Three to five days after the lead letter is sent, an e-mail invitation will be sent to all sample members (Attachment 7). It will again invite the sample member to participate in the survey and will provide a hyperlink to the survey Website.
- One week after the e-mail invitation, a reminder e-mail will be sent to all sample members (Attachment 8). The e-mail will encourage those who have not yet logged in to the Website to participate in the survey.
- One week after the first e-mail reminder, a second e-mail reminder will be sent to all non-respondents (Attachment 8). The e-mail will reinforce the purpose and relevance of the survey.
- One week after the second e-mail reminder, a third e-mail reminder will be sent to all remaining non-respondents (Attachment 8).
- In addition, a final reminder letter (Attachment 9) will be mailed by express mail (FedEx) along with a hardcopy version of the survey. Enclosed with the letter will be a postage paid, business reply envelope for returning the completed questionnaire.

B.3. Methods for Maximizing the Response Rate and Addressing Issues of Nonresponse

The ability to gain the cooperation of potential respondents is key to the success of these two surveys. Consistent with sound survey methodology, the design of the survey will include approaches to maximize response rates, while retaining the voluntary nature of the effort. We will use the following approaches to maximize response rates for the surveys:

- Participation will be made as easy and non-burdensome as possible by designing each questionnaire to take no more than an average of 30 minutes to complete.
- The online instruments will be designed to be clear and easy to understand. Thorough usability testing of the survey instruments will be conducted to eliminate technical errors and to ensure ease of navigation and use.
- Advanced outreach will raise awareness about the surveys and to encourage participation (e.g., announcements on NIH Websites and newsletters).
- The lead letter and introductory e-mail invitations will inform sample members of the study. They will contain enough information to generate interest in the surveys. The letter and email will provide a point of contact at RTI for additional information.

- Follow-up e-mails will remind sample members about the survey, and encourage participation. These reminders will always include a link to the survey.
- A final reminder letter will include a hardcopy version of the survey to provide an alternative mode for answering the questions.

B.4. Tests of Procedures of Methods

The survey instruments have also been tested through a modified Question Appraisal System (QAS). With the QAS, the questions in the instrument were analyzed in relation to the tasks required of the respondents (to understand and respond to the questions) and evaluate the structure and effectiveness of the questionnaire form itself. RTI International's Question Appraisal System (QAS-04) was used to guide this instrument review. This coding system constitutes an item taxonomy that describes the cognitive demands of the questionnaire and documents the question features that are likely to lead to response error. These potential errors include comprehension, task definition, information retrieval, judgment, and response generation. This appraisal analysis was used to identify possible revisions in item wording, response wording, questionnaire formats, and question ordering/instrument flow.

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

Dr. David Wilson
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709
Phone: 919-541-6990
E-mail: dwilson@rti.org

LIST OF ATTACHMENTS

Attachment 1 - Applicant Survey Instrument

Attachment 2 - Reviewer Web-Based Survey Instrument

Attachment 3 – Advisory Council Web-Based Survey Instrument

Attachment 4 - Privacy Act Determination Letter

Attachment 5 - IRB Exemption Letter

Attachment 6 – Lead Letter

Attachment 7 - Email Invitation