

Supporting Statement B for:
Evaluation of the NIH Academic Research Enhancement Award (NIH
OD)

October 2014

Michelle M. Timmerman, Ph.D.
Office of the Director
National Institutes of Health

6705 Rockledge Drive
Bethesda, MD 20892

Telephone: 301.402.0672

Fax: 301.402.0672

Email: michelle.timmerman@mail.nih.gov

Table of contents

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1

B.1	RESPONDENT UNIVERSE AND SAMPLING METHODS.....	1
B.2	PROCEDURES FOR THE COLLECTION OF INFORMATION.....	7
B.3	METHODS TO MAXIMIZE RESPONSE RATES AND DEAL WITH NON-RESPONSE.....	9
B.4	TEST OF PROCEDURES OR METHODS TO BE UNDERTAKEN.....	11
B.5	INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS AND INDIVIDUALS COLLECTING AND/OR ANALYZING DATA.....	12

LIST OF ATTACHMENTS:

1. SURVEY OF AWARDEES WITH SCREENSHOTS
2. SURVEY OF APPLICANTS WITH SCREENSHOTS
3. SURVEY OF STUDENTS WITH SCREENSHOTS
4. SEMI-STRUCTURED INTERVIEW GUIDE
5. PRIVACY IMPACT STATEMENT
6. PRIVACY ACT MEMO
7. IRB APPROVAL LETTERS
8. SURVEY INTRODUCTORY EMAIL
 - 8A. AWARDEES
 - 8B. APPLICANTS
9. SURVEY INTRODUCTORY EMAIL TO AREA STUDENTS
10. SURVEY INVITATION LETTER
 - 10A. AWARDEES
 - 10B APPLICANTS
11. SURVEY INVITATION LETTER TO AREA STUDENTS
12. SURVEY EMAIL REMINDER
 - 12A. AWARDEES
 - 12B APPLICANTS
13. SURVEY EMAIL REMINDER TO AREA STUDENTS
14. SURVEY TELEPHONE FOLLOW-UP SCRIPT
 - 14A. AWARDEES
 - 14B APPLICANTS
15. SURVEY TELEPHONE FOLLOW-UP SCRIPT FOR AREA STUDENTS
16. INTERVIEW INTRODUCTORY E-MAIL
17. INTERVIEW SCHEDULING (PHONE)
18. INTERVIEW REMINDER E-MAIL

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

The AREA program of the NIH provides research grants to meritorious projects at institutions of higher education that receive less than six million dollars of NIH research funding per year. The AREA program not only supports worthy research projects, but aims to enhance the research environment of the institutions receiving AREA funds, and contribute to the United States' scientist workforce by training undergraduate and graduate students in biomedical research. AREA funding across NIH Institutes and Centers (ICs) totaled over 79 million dollars in FY 2012 and more than 77 million in FY 2013. With this much public funding invested annually in the program, it behooves NIH to determine whether or not it is meeting its goals.

B.1 Respondent Universe and Sampling Methods

B.1.1. Respondent Universe

Eligible AREA awardees include all Principal Investigators (PIs) who were awarded AREA (R15) grants from the start of the program in 1985 until the end of 2010. The NIH Query, View, and Report (QVR) database shows that 3,916 R15 awards have been made in this period to a total of 3,288 unique PIs. This group of 3,288 comprises the universe of eligible AREA PIs. The universe of eligible AREA awardees will be sampled to administer the web-based survey (**Attachment 1**). The universe of respondents for the Awardee Semi-Structured Interview (**Attachment 4**) will be AREA PIs who completed the web-based survey. According to our estimates, we expect the latter universe to be the total number of completed surveys from awardees to be 480 (See Table B-1 below).

The population from which the comparison group will be selected is 3,069 unique applicants who never received an award in the history of the program, yet had submitted

projects from 1985 until the end of 2010 deemed worthy of review by a full committee within the appropriate IC. The PIs of proposals that did not make it to full review, and were not scored by peer scientists, are not included in the comparison group universe since their proposals are expected to be significantly inferior to the awardees' proposals in scientific merit. It is expected that unsuccessful applications that made it to the final round of review are more similar to successful applications than those unsuccessful applications that were culled earlier in the review process. The sample of unsuccessful applicants will be surveyed with a web-based applicant survey (**Attachment 2**).

There are two existing sources of information about undergraduate, graduate and health professional students who worked on AREA projects. The first source is administrative supplements to AREA awards to Promote Diversity in Health-Related Research. PIs may apply for supplemental funding to support a student who is a member of an underrepresented minority group, who is disabled, or who is from a disadvantaged background. The supplemental award is used to support the student's work on the AREA-funded research project. There are 151 students named in the Notice of Awards (NoA) supplemental to 129 AREA projects funded 1997-2010.

The second source of information about students is the Final Progress Reports (FPRs) required to be submitted at the end of the project. We extracted 5608 student names from 1024 unstructured reports from 1999 until the end of 2010. Of 4,420 R15 awards made 1985-2010, Final Progress Reports are available for less than 24% (1024 FPRs total).

B.1.2 Sampling Methods for Web Surveys

The survey will be administered to a simple random sample of 600 Principal Investigators who were awarded AREA grants fiscal years 1985-2010. A simple random sample of 600 never successful applicants to the AREA program who applied to the program 1985-2010 will serve as the comparison group.

Semi-structured telephone interviews will be conducted with a total of 50 AREA awardees. A strategic, convenience sample from the group of surveyed awardees will be drawn for the interviews, balanced by institution type and the years the project was undertaken. Interviewing PIs who have already provided basic information about their program participation through the survey, will facilitate efficient use of the semi-structured interview method. Interviewers will not need to spend much time establishing the basic details of the interviewed awardees' experience of the AREA grant, and instead will elicit in-depth exploration of the interviewees' mentoring, institution-capacity building, and research.

A database with the names and any available contact information of students involved in AREA-funded research projects will be constructed from names extracted from the Final Progress Reports. From this database, we will gather a simple random sample of 450 names. Because of the importance of cultivating a diverse scientific workforce to the interests of ensuring that the nation produces the highest quality biomedical research, we will survey the entire population of 151 students who were the beneficiaries of administrative supplements to parent AREA grants.

It is necessary to survey students who participated in AREA-sponsored research since Congress has mandated that the AREA program facilitate the training of the United

States' biomedical scientist workforce. The student surveys will reveal whether or not participation in AREA research has had its intended effect.

Table B-1 shows the target population count, sample size, expected response rate and expected number of completes.

Population	Target Population	Sample Size	Tracing	Expected Response Rate	Expected # of Completes
AREA PIs (surveyed)	3288	600	600	0.80	480
AREA PIs (interviewed)	480	50	50	1.00	50
Applicants (surveyed)	3069	600	600	0.40	240
Diversity Supplement Students	151	151	151	0.50	76
PI-Reported Students	5608	450	450	0.50	225
Overall	8848	1800	1800	0.57	1021

While we expect to have an 80% response rate for former AREA PIs, we assume a 40% response rate from the AREA applicants and a 50% response rate from students who worked on AREA projects. Though different factors may be responsible for non-response, we anticipate two main reasons: (1) failure to locate respondents, and (2) refusal to participate. If respondents cannot be located, there is zero probability of getting them to respond to surveys. Non-respondents in the second group consist of those who are contacted and asked to participate in the survey, but refuse to do so. AREA applicants who were unsuccessful in obtaining AREA funding may not feel connected

enough to the program to respond to the survey. Students may have been previously unaware that their professor was funded by the NIH AREA program and likewise not feel a strong connection to the program. Non-response will be managed in the interview recruitment by selecting a convenience sample. This sample will be strategic so that participants are recruited to represent a diversity of institution types and years of AREA program participation.

Section B.1.4 addresses plans to minimize the first type of nonresponse to surveys, and Section B.3 addresses plans to maximize the survey response rate.

B.1.3. Levels of Precision

The main objective of this study is to compare AREA PIs to unsuccessful applicants in order to assess the degree to which the AREA program has been successful in building the research environment in institutions where the grants have been awarded. These comparisons will be based on survey responses and other program characteristics known from existing NIH databases. The surveys will assist in assessing differences in the number of hands-on research training opportunities available to the student population of AREA-eligible institutions of higher education, the relative quality of these opportunities, and the degree to which AREA grants have facilitated collaboration within PIs' institutions.

The expected sample sizes of 480 PIs and 240 applicants will provide 80% power to detect differences of about 11 percentage points in survey responses, which we expect to be adequate to delineate key differences between these groups.

Another consideration is the proportion of students who go on to pursue research careers. A sample size of 225 students is adequate to estimate this proportion to within 5

percentage points, which is a sufficient level of accuracy for this purpose. If the diversity students are combined with the rest of the PI-identified students, for a sample size of 301, this sample would be adequate to estimate this proportion to within 6 percentage points.

B.1.4 Updating and Discovering Respondent Contact Information

It is necessary to update contact information for awardee and applicant respondents in order to administer the surveys. This information includes e-mail addresses, telephone numbers (work or home) and work or home addresses.

There is little contact information available for students who benefitted from AREA-funded research in either the AREA supplements to Promote Diversity in Health-Related Research or the Final Progress Reports. It is unknown how accurate and useful will be the information reported about their students by PIs on the web surveys. Therefore, it will be necessary to engage in tracing the contact information of all students who comprise the sample.

Westat will employ trained staff to trace contact information (e-mail addresses, employer, home and work addresses, home and work phone numbers) using internet search tools (for example, Google, LinkedIn, People Search) and other methods (such as key term combinations).

B.2 Procedures for the Collection of Information

B2.1 PI and Applicant Surveys

Data collection procedures consist of the following:

- Respondents from the awardee and unsuccessful applicant universes will be selected randomly from the prepared databases via computer.
- The awardees and unsuccessful applicants selected randomly will be traced to update or discover contact information.
- Upon receiving OMB approval, respondents with a valid e-mail address will be sent an introductory e-mail notice about the survey: Survey Introductory E-mail to AREA Awardees and AREA Applicants (**Attachment 8**).

- One week after the e-mail notification, an e-mail invitation letter will be sent to all respondents to encourage participation in the study. The invitation letter requests the individual's participation, introduces the purpose and content of the study, introduces the subcontractor (Westat) and provides contact information for queries. This notification will include the link to the survey (**Attachment 10**).
- One week after the invitation, if a response has not been received, a survey reminder e-mail will be sent to the respondent (**Attachment 12**). If after another one week, no response is received, the reminder e-mail will be re-sent.
- After three weeks of no response from the sampled awardee or applicant, the potential respondent will be called (if a phone number is available) by a professional trained in response conversion, to encourage response to the survey. Up to four (4) attempts will be made to reach the potential respondent by phone (**Attachment 14**).

Data will be gathered using an online survey of approximately 600 AREA PIs (**Attachment 1**) and 600 unsuccessful applicants to the AREA program (**Attachment 2**). The survey will collect information about their research project, publications and other dissemination products, student outcomes, collaboration at the AREA-funded institution, and experiences with NIH and the AREA program. The collection of this data will provide NIH with a better understanding of the AREA program's outcomes and how the program may be improved in the future.

B2.2 Student Surveys

Data collection procedures consist of the following:

- Respondents from the universe of students named in the Progress Reports will be selected randomly via computer.
- 449 students randomly selected and all 151 students identified through AREA administrative supplements will be traced to discover their contact information.
- Respondents with a valid e-mail address will be sent an introductory e-mail notice about the student survey: Survey Introductory E-mail to AREA Students (**Attachment 9**).
- One week after the e-mail notification, an e-mail invitation letter will be sent to all respondents to encourage participation in the study. The invitation letter requests the individual's participation, introduces the purpose and content of the study, introduces the subcontractor (Westat) and provides contact information for queries. This notification includes the link to the survey (**Attachment 11**).

- One week after the invitation, if a response has not been received, a survey reminder e-mail will be sent to the respondent (**Attachment 13**). If after another one week, no response is received, the reminder e-mail will be re-sent.
- After three weeks of no response from the sampled student, the potential respondent will be called (if a phone number is available) by a professional trained in response conversion, to encourage response to the survey. Up to four (4) attempts will be made to reach the potential respondent by phone (**Attachment 15**).

Data regarding student outcomes will be gathered using an online survey of 600 former students who participated in AREA-funded research (**Attachment 3**). The survey will collect information about their participation in AREA-funded research, publications and other dissemination products, educational and career outcomes, and their satisfaction with the research experience. The collection of this data will provide NIH with a better understanding of the impact of the funding on students' careers and how the program may be improved in the future.

B.2.3 PI Semi-Structured Interview

Data collection procedures consist of the following:

- Respondents from the universe of surveyed awardees will be selected strategically by institution type and date of award.
- After OMB approval, respondents will be sent an introductory e-mail notice about the interview (**Attachment 16**). The introductory e-mail notice explains the purpose of the evaluation, requests the awardee's participation, asks for 2-3 dates when he or she may be available for an interview, and asks how the respondent would like to be contacted (that is, by what phone number and e-mail address).
- When a response has been received from PIs contacted about the interview, a phone call will be made by the interviewer with 36 hours to confirm the interview date and time (**Attachment 17**).
- If a response from the PI has not been received after one week of sending the introductory e-mail notice, an experienced telephone interviewer will make up to four attempts for a follow-up call using a script (**Attachment 17**). If a phone number is not available, a reminder e-mail will be sent (**Attachment 18**).
- If the respondent of the phone follow-up or e-mail reminder fails to respond or refuses to participate, another PI with similar characteristics will be selected from the pool of surveyed awardees.

B.2.4 Quality Control

The sub-contractor of this study (Westat) has established and will maintain quality control procedures to ensure standardization and accuracy of data collection and data processing. The contractor will maintain a log of all decisions that affect sample enrollment and data collection. The contractor will monitor response rates and completeness of all acquired data, and provide the OD with progress reports in agreed-upon intervals.

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

B.3.1 Follow-up

To improve response rates, follow-up efforts will be used to encourage survey completion. Efforts to reduce the number of non-respondents consist of the following:

Reminder e-mails will be sent to all non-respondents beginning one week after initiation of the survey (**Attachment 12, Attachment 13**). An identical second reminder will be sent to all non-respondents two weeks after initiation.

Telephone follow-up for non-response will begin three weeks after initiation of the survey. Experienced telephone interviewers, who are trained in non-response conversion, will make up to four attempts for the follow-up calls. The call is a prompting call only encouraging the potential participant to complete the survey. The telephone interviewers will use the script Survey Telephone Follow-up Script for AREA PIs and Applicants (**Attachment 14**) and Students (**Attachment 15**) for the call.

B.3.2 Unit Nonresponse

As discussed above, to improve nonresponse, experienced interviewers trained in nonresponse conversion will re-contact non-responders. However, individuals who make a hard refusal, that is, request not to be called again, will not be contacted. To deal with unit nonresponse due to hard refusals, weights will be used to adjust for nonresponse within cells identified to key variables known (for example, hard refusers' application years).

B.3.3 Item Nonresponse

Although procedures are designed to maximize item response rates, the analysis will need to confront the issue of missing data. Experience with previous surveys indicates that some respondents will omit responses to specific items, particularly sensitive items, although they may have provided most of the data requested. By employing good survey data collection practices, the amount of missing data on any single variable will be minimized to a very low level. However, if item nonresponse is unexpectedly high for any of the key analytic variables, hot deck imputation techniques will be used to estimate missing-item values.

For analyses involving just one or two variables that have not been subject to imputation, we will omit cases with missing data; or, in the case of categorical response variables, we will use an explicit "missing" or "unknown" category. When multivariate techniques involving several variables are used, analytic techniques for missing values will be employed, such as using the variable mean or adding a dummy variable to reflect how the non-respondents differ from other individuals.)

B.4 Test of Procedures or Methods Undertaken

During the 60-day Federal Register Notice (FRN), Westat conducted a pretest of each drafted survey instrument (**Attachments 1, 2, and 3**). The purpose of these pretests was to refine the collection of information, improve clarity and ease of understanding, minimize burden and improve utility.

Convenience sampling was used to select respondents for a pretest of the survey. The pretesting took place in May 2014. Two (2) former AREA awardees, two (2) AREA applicants, and two (2) students who had worked on AREA grants were notified by e-mail of the AREA evaluation and asked to participate in the survey pretest. Upon consent from the pretest respondents, surveys were mailed to the respondents overnight. Respondents completed a paper version of the survey and mailed it back to Westat overnight in a pre-addressed, stamped envelope. Respondents then participated in a follow-up telephone interview that captured their experience with the survey, including the time needed to complete, questions or instructions that were confusing, lists of response options that seemed incomplete, and information that was difficult to provide. Interviews generally lasted about 60 minutes and explored key survey questions. Probes specific to individual responses were also developed based on a review of the returned surveys. Findings from the interview guide pilot test were summarized and necessary changes to the interview guide were made. Each interviewee was provided with \$100 by check as appreciation for their time and effort.

B.4.1 Key Findings and Survey Modifications

This section summarizes key findings from the pretest and modifications made to the surveys in response to the findings, where applicable.

B.4.1.1 Awardee Survey

Completion time: Respondents varied greatly in the time to complete: 120 minutes and 20 minutes. Modifications since made to the survey have significantly shortened its length: a question asking respondents to list names and contact information for all students who worked on their AREA project has been eliminated. A question that asked for the role of each co-author on AREA-funded publications has been eliminated also.

Overall feedback: Respondents were willing to be thorough on the survey since they felt they benefited greatly from the AREA program.

Question 6: Respondents thought the original wording of the question was unclear, so more explanation was added.

Question 9: The following questions ask about future research. Question 9 was added (“Are you still active as a researcher?”) to allow for respondents who are not conducting research to skip the following questions.

Question 12: One respondent was retired and he did not see his circumstances adequately represented in the available choices. A response for retiring researchers was added.

Questions 13-15: Respondents found the wording of the original questions confusing. Wording was revised to ask about: students who worked in the awardee’s lab (Q13), students who worked on the awardee’s AREA project (Q14) and students who disseminated results (Q15).

Questions 22-27: One respondent thought collaboration with researchers outside of awardees’ institution was a marker of research success. Therefore an option to indicate collaborators outside of awardees’ institution was added.

B.4.1.2 Applicant Survey

Completion time: Respondents indicated that the survey took an average of less than 20 minutes to complete.

Overall feedback: Respondents believed the survey would be helpful if the NIH implements changes to the application process based on responses. One respondent said that applicants are not going to go back through 8-10 years of old records to answer these questions. “Don’t remember” options were added to provide an appropriate choice for information that might not be recalled.

Question 5: This question originally did not include a “don’t recall” option. It was added when one respondent said he was uncomfortable checking N/A when he did not remember.

Questions 10 and 11: Changed the DK/NA option to “don’t remember” since one respondent said “don’t remember” is a more accurate description of why he could not provide a number of students.

Question 18: The original wording was unclear and respondents said they had to read it several times to understand it. Question was changed to “Did you remain active as a researcher after submitting your AREA application?”

B.4.1.3 Student Survey

Completion time: Former AREA students reported that the survey took an average of 16 minutes to complete.

Overall feedback: Respondents did not report needing to look up any information to complete the survey.

Question 7: One respondent took an independent research class and chose a professor to work with after enrolling. The other had applied for, and been admitted to, a special undergraduate research program. Neither saw their experiences represented in the choices. “Applied to special research program” and “Took independent research class” options were added.

Question 9: An additional choice was added “Collected specimen samples or data from human subjects,” since one respondent said most of her time in the lab was spent collecting human specimen samples.

Question 16: This question was added to allow respondents who did not disseminate findings from their AREA-funded research opportunities to skip Question 17 requesting information about publications. Both respondents had not disseminated any findings from their experiences with the AREA program.

Question 19: This question originally asked for the frequency that the respondent engaged in activities (i.e.: never, once, two to three times, etc.). Respondents had difficulty remembering the number of times they had engaged in these activities. It was decided that, for analysis, a simple binary (yes/no) will suffice. “Broadly defined” was added to the description “biomedical and behavioral sciences” since respondents felt that both of their fields (bioengineering and dentistry) were only broadly biomedical.

Question 21: Options for post docs, interns, practicum students and residents were added to the list of kinds of trainees and students at the suggestion of both respondents.

The final versions of all three survey instruments and the semi-structured interview guide have been submitted to the Office of Management and Budget for the 30-day FRN.

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

The following individuals were critical in developing the research plan, the conceptual framework, survey questions, and sampling strategies underlying evaluation of the AREA program. Many of the same individuals will be involved with analysis once the data are collected.

Jocelyn Marrow, Ph.D. Project Director Westat 1600 Research Boulevard Rockville MD 20850 240-314-5887 jocelynmarrow@westat.com	Sophia Tsakraklides, Ph.D. Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 301-738-3580 sophiatsakraklides@westat.com
Atsushi Miyaoka Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 301-610-4948 MIYAOKA1@westat.com	Martha Palan Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 MarthaPalan@westat.com
James Bethel, Ph.D. Senior Statistician Westat 1600 Research Boulevard Rockville MD 20850 301-294-2067 BETHELJ1@westat.com	Rene Gonin, Ph.D. Senior Statistician Westat 1600 Research Boulevard Rockville MD 20850 301-517-8084 GONINR1@westat.com
Jennifer Crafts, Ph.D. Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 301-610-4881 CRAFTSJ@westat.com	Kerry Levin, Ph.D. Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 301-738-3563 LEVINK@westat.com
Jocelyn Newsome, Ph.D. Subject Matter Expert Westat 1600 Research Boulevard Rockville MD 20850 301-212-3734 JocelynNewsome@westat.com	