Supporting Statement B

GuLF Study:

Gulf Long-Term Follow-Up Study for Oil Spill Clean-Up Workers and Volunteers

(NIH/NIEHS)

Submitted: 1 December 2010

Revised: 20 January 2011

Dale P Sandler, PhD
Principal Investigator
Chief, Epidemiology Branch
National Institute of Environmental Health Sciences
PO Box 12233
Research Triangle Park NC 27709
Phone: 919-541-4668
Fax: 919-541-2511
Email: sandler@niehs.nih.gov

Richard K Kwok, PhD
Lead Associate Investigator
PO Box 12233
Research Triangle Park NC 27709
Phone: 919-627-8892
Fax: 919-541-2511
Email: richard.kwok@nih.gov

TABLE OF CONTENTS

LIST OF ATTACHMENTS & APPENDICES

Materials translated into Spanish and Vietnamese will be submitted to OMB as a non-substantive change at a later date.

COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

B.1. Respondent Universe and Sampling Methods

There are 150,000 names on the PEC (worker training) list but we currently have no contact information for 40,000 leaving us with 110,000 potential participants. We expect an additional 20,000 names to be identified from non-overlapping persons on the NIOSH list, and smaller groups of workers from Parishes and some federal groups (e.g. Fish and Wildlife Service and the Coast Guard) that did not complete a training module via PEC. In total, we will have 130,000 names. Eliminating individuals represented more than once (because they took a second training module, for example) and those under age 21, we expect a total of 110,000 potential participants. We anticipate needing to attempt to reach 86,000 in order to enroll a cohort of 55,000 workers (a 60-65% participation rate) who complete the enrollment questionnaire. These 55,000 persons will comprise the full cohort. Among this group, we estimate that about 43,000 (80%) will have engaged in clean-up activities while the remaining 12,000 (~20%) did not. These 12,000 unexposed persons will include up to several thousand Federal responders who engaged only in response activities such as administrative, oversight, or logistical support that did not involve any contact with spill-related oil, oil byproducts, or dispersants.

We plan to select a probability sample of workers for contact from the universe of potential participants, including Federal and other workers obtained from databases not yet received. To do this, we will field the sample in replicates that are randomly selected from the universe of potential participants. Workers in late arriving databases will be randomly selected and added to each replicate so that they have the same probabilities of selection as the original replicate members. Stratified random sampling will be used to select workers for these replicates, using sampling probabilities that will be adjusted and recorded over time to account of greater (or lower) than expected response rates and to make course corrections if the assumptions about the likelihood of exposure in a stratum is incorrect. Possible stratification variables include:

- Date trained: an indicator of length and possibly level of exposure--before July 1, 2010 versus after July 1, 2010

- Badge-in location: an indicator of where they worked – off shore (closer to the source or burning oil), on land (e.g., beach clean-up), or trailers (office) locations.

We will then assign sampling probabilities to strata based on the expected distribution of exposed workers on the cohort. For example, we may initially select 100% of those who trained early and badged in for off shore locations but a smaller fraction of those who trained later and badged in for office or beach clean-up locations in order to achieve a cohort that represents the range of clean-up activities and includes sufficient numbers of the potentially more highly exposed. We will reassess the sampling fractions periodically in deciding the size and composition for replicates yet to be fielded for telephone interviews. For instance, we will reassess the sampling fractions after attempting to contact the first 2,000 names and after batches of 5,000 names thereafter. We will increase or decrease the selection probabilities used for each stratum to account for differences between the expected and actual response rates and the distribution of potential exposures within strata.

Because it is important to complete home visits shortly after completing the telephone interview, it will not be possible to fully characterize the telephone interview respondents in order to select the sample for the Active Follow-up Sub-cohort. Thus, we will employ a two-stage design which is based on sampling probabilities preset and recorded on the data records for each sample replicate which will be compared to a random number that is generated and also recorded on the record. After excluding workers residing outside the 4 Gulf states, the remaining workers will be selected for the Active Follow-up Sub-cohort when the random number is less than the probability of selection. These selection probabilities may be modified as we learn more about response rates and work experience. The sampling probabilities used for each telephone respondent will vary across strata defined based on responses to questions such as work experience, proximity of their residence to the Gulf (75% of workers in the Gulf states are from communities proximate to the Gulf). The self-reported main tasks will trigger a computer algorithm that will tell the interviewer if a participant is to be recruited for the active follow-up sub-cohort. We will enroll as many of the most highly exposed as possible. We will invite those from out of state to join if we find that persons in the potentially most highly exposed tasks are under-represented. These assumptions are represented in the burden request.

Groups of workers such as HAZWOPERs, those reporting work at the source, on the oil rig or those on boats near burning oil are expected to be selected with 100% probability, but workers with office only jobs or beach-clean-up jobs (where the N is substantially greater than that for workers at source control) may be selected with probabilities closer to 50%, in order to achieve a subcohort for active follow-up of 24,000 individuals among whom at least 6,000 are unexposed. Again, we will evaluate our assumptions about yields and response rates periodically, say after 2,000 telephone interviews and then periodically at 5,000 interview increments, to determine if adjustments are needed to the sampling probabilities being used for each stratum. Ideally, we would not need to make many adjustments to the sampling fractions, but the analysis will take into account variations in sampling probabilities used across replicates. Because each potential participant is randomly assigned to a replicate for the telephone interview and then randomly selected for active follow-up if responding, each sampled worker derived from the cross of the two strata (telephone interview and active follow-up) will have identical unconditional probabilities of selection, even if unequal selection probabilities are used across replicates.

Many investigators have published on two-stage designs. Our approach is adapted from a method called randomized recruitment, first described by Weinberg and Sandler in 1991 [Weinberg CR, Sandler DP: Randomized recruitment in case-control studies. Am J Epidemiol 134:421-432, 1991] and used in several studies at NIEHS and many others elsewhere since then. Although the approach was developed in the context of case-control studies, the approach has been adapted for other designs.

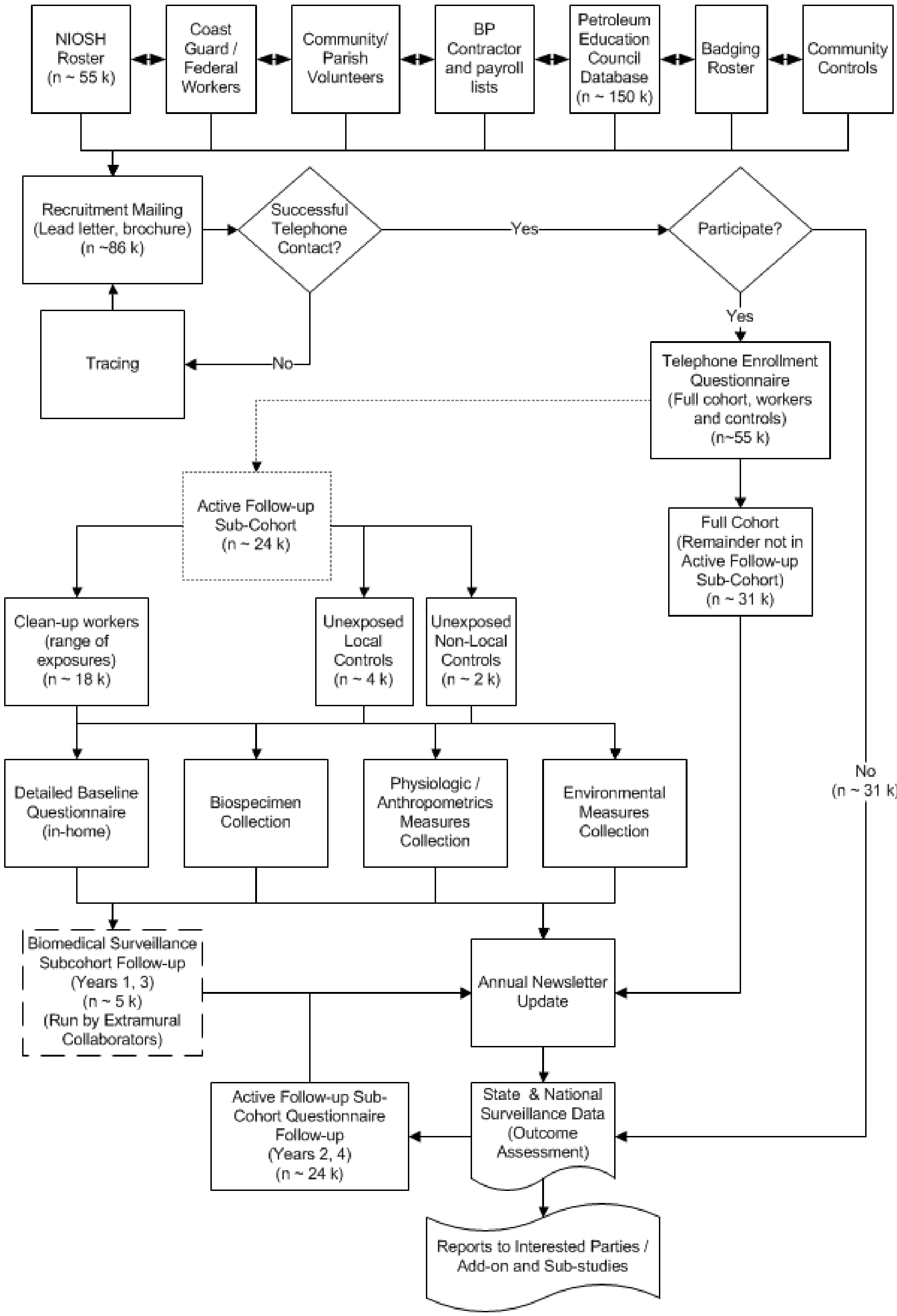
There are sufficient eligible persons to recruit 18,000 workers and 6,000 controls into the Active Follow-up Sub-cohort, assuming a 70% participation rate among persons who have already enrolled in the full cohort by participating in the telephone interview. The size of the Active Follow-up Sub-cohort has been capped at 24,000 in light of available funding and statistical power considerations; the base population is large enough that this target is achievable even with a modestly lower participation rate. Based on current information, we estimate that about 26% of the eligible controls are from outside the immediately affected communities. By oversampling these non-local controls, we expect to recruit approximately 2,000 non-local controls and 4,000 local controls, with both groups including Federal controls as described above.

The expected participation rates provided above are reasonable, given anecdotal reports from collaborating federal agencies, media reports, and feedback from community groups and focus groups of clean-up workers that indicate widespread concern about potential health effects from the oil spill among clean-up workers and members of the affected communities. Furthermore, it is possible that the eventual cumulative total of workers will be greater than is currently estimated. We will know the real total only after we have obtained worker lists from other agencies and local communities engaged in clean-up and crossed the lists to identify unique additional workers who did not complete PEC training. We anticipate needing to modify sampling fractions (and tracking them) as we learn more about the cohort – both through information we collect and information we are still negotiating with BP and others to receive.

In any case, power calculations indicate that even if actual participation rates turn out to be as much as 20% lower than those indicated above, this study will still be sufficiently powered to achieve its specified aims, with an increase in minimum detectable ORs or differences of less than 10-15%.

About 18,000 exposed workers and 6,000 controls will be Active Follow-up Sub-cohort participants. The rest of the full cohort (N~31,000) will comprise individuals to be passively followed who either were not randomly sampled to be part of the Active Follow-up Sub-cohort or who refused to be part of the Active Follow-up Sub-cohort (but participated in the enrollment telephone interview).

Schematic of Study Design



Thus, the total size of the full cohort is anticipated to be approximately 55,000 persons (43,000 workers and 12,000 controls), consisting of 24,000 members of the Active Follow-up Sub-cohort (18,000 workers and 6,000 controls [4,000 local and 2,000 non-local, including Federal]) and 31,000 passively followed members of the full cohort (25,000 workers and 6,000 controls).

Based on other prospective observational studies, we anticipate 90% follow-up and participation in telephone interviews after enrollment for the Active Follow-up Sub-cohort. Thus, completed follow-up interviews are expected for approximately 18,000 workers and 5,400 controls in Years 2 and 4.

This study is designed not around a few narrow *a priori* hypotheses, but rather to allow the investigation of a wide range of potential adverse health effects. The study size and the number of individuals who experienced a given exposure – and the consequent statistical power – have largely been determined by the number of individuals involved in the clean-up operations and their distribution by task/exposure. While this study will have limited power to examine certain rarer exposures or outcomes in the near future, this is the largest study to date of oil spill clean-up workers and it is important that we address, to the extent feasible, the wide range of public health concerns. It is a prospective study and as time passes, if the exposure continues to exert an impact on some health outcomes, power will increase.

Table 3 presents minimum detectable odds ratios across a range of proportions of exposure among the workers and of health outcome among the controls. Estimates are shown separately for analyses of the full cohort and of the Active Follow-up Sub-cohort, including all controls or including only the non-local controls. Estimates are also shown for analyses of the Biomedical Surveillance Sub-cohort. All estimates are based on a two-sided test with α=5% and power=80%. As the table shows, this study has excellent power to detect small risks, except when exposure or outcome is rare. For example, in an analysis of the full cohort, if 10% of the workers received a given exposure (e.g., high exposure to VOCs) and the incidence or prevalence of disease is 1%, this study would have sufficient power to detect an OR of at least 1.56 when using all 12,000 controls and 1.86 when using only the 2,500 non-local controls. In an analysis restricted to the Active Follow-up Sub-cohort, with proportion of exposure of 10% and disease incidence/prevalence of 10%, the minimum detectable OR would be only 1.27-1.32

Table 3. Minimum detectable odds ratios for a range of proportions of exposure among the workers and for all controls vs. non-local controls, based on a two-sided test with α=5% and power=80%

Size of control group (i.e., all vs. non-local)	Proportion (N) of workers exposed to a given agent					
	5%	**10%**	**25%**	**50%**	**75%**	**100%**
Full cohort: 43,000 workers, 12,000 controls:						
	N=2,150	N=4,300	N=10,750	N=21,500	N=32,250	N=43,000
Proportion of controls with outcome=1%						
12,000[a]	1.74	1.56	1.41	1.35	1.33	1.32
2,500[b]	2.02	1.86	1.76	1.72	1.71	1.70
Proportion of controls with outcome=10%						
12,000[a]	1.23	1.17	1.13	1.11	1.10	1.10
2,500[b]	1.30	1.25	1.22	1.21	1.21	1.21
Proportion of controls with outcome=30%						
12,000[a]	1.15	1.11	1.08	1.07	1.07	1.07
2,500[b]	1.19	1.16	1.14	1.14	1.14	1.13
Active Follow-up Sub-cohort: 18,000 workers, 6,000 controls:						
	N=1,000	N=2,000	N=5,000	N=10,000	N=15,000	N=18,000
Proportion of controls with outcome=1%						
6,000[a]	2.20	1.88	1.63	1.53	1.49	1.47
2,000[b]	2.38	2.12	1.93	1.86	1.83	1.82
Proportion of controls with outcome=10%						
6,000[a]	1.36	1.27	1.19	1.16	1.15	1.15
2,000[b]	1.40	1.32	1.27	1.25	1.24	1.24
Proportion of controls with outcome=30%						
6,000[a]	1.24	1.18	1.13	1.11	1.10	1.09
2,000[b]	1.26	1.21	1.17	1.16	1.16	1.15
Biomedical Surveillance Sub-cohort: 4,500 workers, 500 controls:						
	N=250	N=500	N=1,250	N=2,500	N=3,750	N=5,000
Proportion of controls with outcome=1%						
500[a]	4.48	3.78	3.28	3.09	3.02	2.99
Proportion of controls with outcome=10%						
500[a]	1.89	1.71	1.58	1.54	1.52	1.51
Proportion of controls with outcome=30%						
500[a]	1.58	1.46	1.37	1.34	1.33	1.33
Full cohort: 43,000 workers, 12,000 controls:						
	N=2,150	N=4,300	N=10,750	N=21,500	N=32,250	N=43,000
Proportion of controls with outcome=1%						
12,000[a]	1.74	1.56	1.41	1.35	1.33	1.32
3,300[b]	1.94	1.78	1.66	1.62	1.61	1.60
2,500[c]	2.02	1.86	1.76	1.72	1.71	1.70
Proportion of controls with outcome=10%						
12,000[a]	1.23	1.17	1.13	1.11	1.10	1.10
3,300[b]	1.28	1.23	1.20	1.19	1.18	1.18
2,500[c]	1.30	1.25	1.22	1.21	1.21	1.21
Proportion of controls with outcome=30%						
12,000[a]	1.15	1.11	1.08	1.07	1.07	1.07
3,300[b]	1.18	1.15	1.13	1.12	1.12	1.12
2,500[c]	1.19	1.16	1.14	1.14	1.14	1.13
Active Follow-up Sub-cohort: 20,000 workers, 7,000 controls:						
	N=1,000	N=2,000	N=5,000	N=10,000	N=15,000	N=20,000
Proportion of controls with outcome=1%						
7,000[a]	2.12	1.82	1.59	1.50	1.46	1.44
2,800[b]	2.28	2.00	1.80	1.72	1.70	1.68
2,000[c]	2.38	2.12	1.93	1.86	1.83	1.82
Proportion of controls with outcome=10%						
7,000[a]	1.34	1.25	1.18	1.15	1.14	1.14
2,800[b]	1.38	1.30	1.24	1.21	1.21	1.20
2,000[c]	1.40	1.32	1.27	1.25	1.24	1.24
Proportion of controls with outcome=30%						
7,000[a]	1.22	1.17	1.12	1.10	1.09	1.09
2,800[b]	1.25	1.19	1.15	1.14	1.13	1.13
2,000[c]	1.26	1.21	1.17	1.16	1.16	1.15
Biomedical Surveillance Sub-cohort: 4,500 workers, 500 controls:[separate clinical exemption]						
	N=250	N=500	N=1,250	N=2,500	N=3,750	N=5,000
Proportion of controls with outcome=1%						
500[a]	4.48	3.78	3.28	3.09	3.02	2.99
Proportion of controls with outcome=10%						
500[a]	1.89	1.71	1.58	1.54	1.52	1.51
Proportion of controls with outcome=30%						
500[a]	1.58	1.46	1.37	1.34	1.33	1.33

[a] All controls in cohort/sub-cohort

[b] Non-local controls in cohort/sub-cohort

when using the full control group (N=6,000) or the non-local control group (N=2,000). The Biomedical Surveillance Sub-cohort, with 4,500 workers and 500 controls, provides adequate statistical power to detect odds ratios of at least 1.58 when 25% of workers received a given exposure and the incidence or prevalence of disease is 10%. For perspective, estimated relative risks of lower respiratory tract symptoms observed among clean-up workers in previous oil spills ranged from 1.5 to 3.6 [Janjua, et al. 2006, Zock, et al. 2007, Meo, et al. 2009, Sim, et al. 2010]. Thus GuLF Study is sufficiently powered to observe such relative risks for these outcomes.

Minimum detectable differences for continuous outcomes are presented in Table 4.

Table 4. Minimum detectable differences, in standard deviations, for continuous outcomes for a range of proportions of exposure among the workers and for all controls vs. non-local controls, based on a two-sided test with α=5% and power=80%

Size of control group (full vs. non-local)	Proportion of workers exposed to a given agent					
	5%	**10%**	**25%**	**50%**	**75%**	**100%**
Full cohort: 43,000 workers, 12,000 controls:						
	N=2,150	N=4,300	N=10,750	N=21,500	N=32,250	N=43,000
12,000[a]	0.066	0.050	0.037	0.032	0.030	0.029
2,500[b]	0.082	0.071	0.062	0.059	0.058	0.058
Active Follow-up Sub-cohort: 18,000 workers, 6,000 controls:						
	N=1,000	N=2,000	N=5,000	N=10,000	N=15,000	N=18,000
6,000[a]	0.101	0.075	0.055	0.046	0.043	0.041
2,000[b]	0.109	0.089	0.074	0.069	0.067	0.066
Biomedical Surveillance Sub-cohort: 4,500 workers, 500 controls:						
	N=250	N=500	N=1,250	N=2,500	N=3,750	N=5,000
500[a]	0.217	0.177	0.148	0.137	0.133	0.131

[a] All controls in cohort/sub-cohort
[b] Non-local controls in cohort/sub-cohort

Differences are expressed in standard deviations (SDs) and are based on a two-sided test with α=5% and power=80%. Results are shown separately for analyses of the full cohort and of the Active Follow-up Sub-cohort including all controls or including only the non-local controls. In addition, estimates are shown for analyses of the Biomedical Surveillance Sub-cohort. This table demonstrates that the present study has sufficient power to detect small differences in continuous outcomes. For example, in an analysis of the full cohort that examines an exposure of 10% prevalence, we will be able to detect minimum differences of less than 0.050-0.071 SD. A similar analysis in the Active Follow-up Sub-cohort will be able to detect minimum differences of less than 0.09 SD (0.075 when using all 6,000 controls and 0.089 when using the 2,000 non-local controls). Such an analysis in the Biomedical Surveillance Sub-cohort will have sufficient power to detect a minimum difference of 0.177 SD. For perspective, in a study of volunteers involved in the Prestige oil spill clean-up and unexposed controls [Laffon, et al. 2006], results of the comet assay in peripheral blood leukocytes showed differences between the two groups of approximately 4.3 SD in comet tail length. A study of health effects related to the Tasman Spirit oil spill found a difference of about 0.6 SD in symptom scores between coastal residents affected by the spill and persons living away from the site of the spill [Janjua, et al. 2006]. The present study is very well powered to detect such effects. Finally, power calculations indicate that even if participation rates turn out to be as much as 20% lower than expected, the minimum detectable ORs or differences will increase by less than 10-15%.

B.2. Procedures for the Collection of Information

This is a non-probability sample and represents a subset of the population with unique exposures. To capture a representative sample of the clean-up workers and controls, we will target individuals across the various categories of job/potential exposure from the Petroleum Education Council (PEC), National Institute of Occupational Safety and Health (NIOSH), or other worker/volunteer rosters and administrative lists. These individuals are potential participants because they are believed to have engaged in clean-up work or participated in worker training modules in anticipation of such work. We will exclude individuals such as journalists who did not engage in clean-up activities but were required to undergo safety training to gain access to worker staging areas (and, therefore, may appear on the PEC list). These individuals will be determined from either the training lists (i.e., individuals who indicated that they intended to work for less than one week) or via screening questions during the enrollment telephone interview. We will use data from our planned mini-pilot (at the beginning of field work) to determine the feasibility of also *efficiently* identifying and excluding individuals such as caterers and administrative/office staff who engaged in clean-up *related* activities, but not clean-up activities *per se*; however, this issue is complex and requires data that will become available only after we go into the field. We define potentially *exposed* subjects as individuals who completed at least one day of oil-spill clean-up-related work, either paid or volunteer. We define *unexposed* subjects as eligible individuals who completed safety training in anticipation of performing clean-up work but did not do so. Additional unexposed subjects will be recruited from among Federal workers eligible for clean-up work but ultimately not deployed. Selection for the Active Follow-up Sub-cohort will cover all levels of potential exposure but will oversample workers with the highest potential exposures to oil, oil byproducts and dispersants.

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

While we have 130,000 names, there will be a large percentage with information insufficient to contact them. We have been working with BP to address this. The two largest cuts will be those with no information other than name (~30,000) and those < 21. We will further exclude any for whom we have evidence that they took training but did not intend to do oil-spill clean up work (reported on the training form). There will be a small fraction of those we contact who will be found ineligible because they neither worked nor intended to work in any capacity on clean-up. This includes the press, government officials and others who passed through the site as observers but needed to badge in. If response rates are unexpectedly high we plan to establish a sampling algorithm to ensure adequate representation of all relevant tasks and locations as well as unexposed individuals – aiming for a cohort of approximately 80% with some clean-up specific job and 20% without. Unfortunately, selection criteria will need be adjusted as we go (see below) due to the need to get into homes as soon as possible after phone interview and the many unknowns even at this stage.

Response bias will be a greater issue in cross-sectional analyses than for prospective analyses which are the main purpose of the study. Some subset of the population were rostered by NIOSH. For those we will have information on intended job, demographics, smoking and a few other useful variables that will allow us to compare responders and non-responders. We are planning to collect some basic information (simplified job information, demographics) by phone from proxy respondents when workers are reported to be deceased or too ill to participate. We will have the full range of telephone interview data to compare responders and non-responders for the home visit stage.

There is little we can do to address differences between those with sufficient contact information and those without. Depending on the level of information we are able to get from BP and its contractors, we should be able to compare these groups by age, type of training, place of training, and purpose of training. We will have additional information to compare responders and non-responders to the telephone survey, including age, gender, language spoken (for some) and for those listed on the NIOSH roster, additional factors. We will have extensive occupational and demographic information from the telephone interview to compare those who do and no not agree to participate in the home visit. In addition, we will be able to assess differences in mortality and cancer incidence in those who do and do not complete the home interview because both groups will be followed over time through death certificate (National Death Index) and cancer Registry linkages. Our analysis of response bias will include simple comparison of frequencies of specific characteristics as well as more comprehensive assessment of bias by exploring the impact of selection factors on specific health outcomes.

Home interviews scheduled at the convenience of the participant will maximize response rates. Participants will be sent a reminder about the appointment for the interview and the importance of completing the requirements of the study. Non-responders will be sent follow-up reminders by mail, and are subsequently contacted by phone to determine whether or not they wish to continue their participation. We will conduct interviews in English, Spanish, and Vietnamese. Special accommodation will be made for those speaking other languages (e.g. Haitian Creole, etc.), if feasible and warranted by the number of workers speaking these languages. PEC training was conducted in English, Spanish, and Vietnamese only so we do not anticipate a large number of those speaking other languages. However, should this change based on data from the PEC list or input from community groups, we will submit an amendment to the IRB with appropriate translated documents for approval.

B.4. Test of Procedures or Methods to be Undertaken

We established and continue to solicit new contacts with several community organizations, representative worker organizations, advocacy groups, and state and local government representatives to identify the primary health issues of concern locally and to discuss study implementation issues across the five state area.

We have conducted a series of meetings with state and local health department representatives as well as with the NGOs that span the various advocacy and occupational groups representing the workers involved in clean-up throughout the Gulf. We met with groups in Mississippi and Alabama during the week of September 12, 2010; Florida the week of September 19, 2010; and Louisiana during the week of October 3, 2010. Although we have been in contact by telephone and email, we are currently working to schedule in-person meetings in Texas. The groups we have contacted span cultural, religious, occupational, and state and local government sectors and are continuously updated as more information and contacts are made (current as of 10/22/2010).

All procedures and questionnaires underwent internal testing prior to implementation, and are modeled on proven methods in previous studies. Finally, the information gleaned from each activity allows further refinement of all study materials and procedures. Forms were shortened and modified to streamline data collection, thus reducing the burden on participants.

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

Dr. Dale P Sandler, Chief, Epidemiology Branch, NIEHS, the Principal Investigator on this study, developed the statistical approach for the study in conjunction with the study team listed below. Data will be collected and managed by SRA. Data will be analyzed by the study team listed.

Richard K Kwok, PhD, Lead Associate Investigator, NIEHS (919) 316-4803
Lawrence Engel, PhD, Associate Investigator, Memorial Sloan-Kettering and NIEHS (646) 735-8171
Stephanie London, MD DrPH, Associate Investigator, NIEHS (919) 541-5772)
Aubrey Miller, MD MPH Associate Investigator, NIEHS (301) 496-3511
Christine Parks, PhD, Associate Investigator, NIEHS (919)541-2577)
Aaron Blair, PhD, Consultant, NCI (301) 496-9094
John Hankinson, PhD, Consultant, Hankinson Associates info@hankconsulting.com
Mark Stenzel, Consultant, Exposure Assessment Applications, LLC (703) 532-2755
Patricia A Stewart, PhD, Consultant, Stewart Exposure Assessments, LLC (703) 532-2755