Background and Strategy for Exposure Assessment for the "Gulf Long-Term Follow-Up Study"

Patricia A. Stewart, Ph.D., C.I.H.
Consultant
Stewart Exposure Assessments, LLC.

Mark Stenzel, M.S., C.I.H.
Consultant
Exposure Assessment Applications, LLC.

Richard K. Kwok, Ph.D.
Lead Associate Investigator
Epidemiology Branch
Division of Intramural Research
National Institute of Environmental Health Sciences

Aaron Blair, Ph.D.
Consultant
Occupational and Environmental Epidemiology Branch
National Cancer Institute

Lawrence S. Engel, Ph.D.
Associate Investigator
Epidemiology Branch
Division of Intramural Research
National Institute of Environmental Health Sciences

Dale P. Sandler, Ph.D.
Principal Investigator
Division of Intramural Research
Epidemiology Branch
National Institute of Environmental Health Sciences

February 26, 2011

INTRODUCTION

On April 20, 2010, the Deepwater Horizon oil drilling rig exploded, killing 11 people and resulting in the largest oil spill in the history of the United States. Within a few weeks, the oil reached land, covering not only the beaches and marshes, but also birds and other wildlife. Thousands of people volunteered and were recruited to remediate the effects of the disaster, while BP and its contractors attempted to stop the well. By July 15, oil was no longer leaking from the well and on September 19, the well was permanently stopped. Even today, however, workers continue to be involved in clean-up activities.

Almost since the beginning of the spill, clean-up workers have been reporting symptoms and ill health effects (NIOSH, 2010a and b), but it is not known if clean-up exposures or activities are responsible. A number of chemical agents were present during the clean-up activities, including volatile organic compounds (VOCs) such as benzene, toluene, xylene, and ethyl benzene (BTEX); semi-volatile compounds, such as polycyclic aromatic hydrocarbons (PAHs) and higher molecular weight alkanes and aromatic hydrocarbons; 2-butoxyethanol (2-BE); propylene glycol (PG); particulate matter (PM2.5 and PM 10.0); oil mist; carbon monoxide; diesel and gasoline engine exhaust; heat; noise; and stress. In addition, workers often worked long hours, may have gone weeks without a day off, and had substantial sun and heat exposure. Musculoskeletal injuries and cuts also have been reported.

The National Institute for Environmental Health Sciences (NIEHS) has initiated an epidemiologic study to investigate the health of the clean-up workers (GuLF Study) (Sandler et al., 2010). A cohort of approximately 55,000 workers will be assembled for a long-term follow-up study. The entire cohort will be administered a questionnaire to obtain information on the subjects' health and the clean-up jobs in which they were engaged. A sub cohort (the Active Follow-up Sub-cohort) will be visited in the home for administration of an additional questionnaire and for the collection of blood, urine, hair, fingernails, and household dust and anthropometric and physiologic measurements. A smaller sub-cohort (the Biomedical Surveillance Sub-cohort) will be asked to also participate in more detailed neurological and physiologic measurements.

A critical component of the GuLF Study will be to characterize possible worker exposure to a number of chemical and physical agents associated with crude oil, dispersants, and other chemicals arising from the spill or used in the clean-up work. Reliable and valid exposure assessments are essential to evaluate and accurately characterize exposure and disease relationships (Hill, 1965). Exposure assessment is challenging because in many epidemiologic studies the exposure data are less than complete.

A typical approach for assessing exposures in cohort studies of workers involves visits by study investigators to the worksite to take worker and environmental measurements; to interview workers, to collect information on sources of exposure, engineering controls for those exposures, tasks (a component of a job), personal protective equipment (PPE), and changes over time; and to collect historical records covering similar topics. Using this information, exposure scenarios are developed for unique combinations of job tasks, work location, personal characteristics, and time. One or more scenarios are assigned to each study subject, based on information from questionnaires or records, to represent the exposure experience(s) of the subject. For example, a worker burning oil on the surface of the water could be assigned a different exposure scenario than a worker burning oil using a flare on the rig ships. Exposure scenarios have been called elsewhere homogeneous or similar exposure groups. Exposure distributions are discussed in more detail in Appendix A.

Depending upon the amount of measurement information available, an exposure estimate, or range of estimates, is developed for each exposure scenario. Exposure estimation methods include calculation of means and other metrics from measurements; statistical modeling of measurements; mathematical modeling using chemical and physical properties of the agent of interest (Jayjock *et al*., 2009); determinant modeling, using measurements and determinants of exposure (e.g., distance from the wellhead or temperature); and professional judgment (see Exposure reconstruction**,** Modeling of Inhalation Exposure Levels) (Viet *et al*., 2008). The estimates are then assigned to each study subject through the exposure scenario link. Thus, all individuals with the same reported information get the same exposure estimate, which is often called a job exposure matrix (JEM). Assessing exposure levels of individuals is not usually possible, because of the lack of measurement data on most individuals. Finally, where possible, the estimates are then evaluated for reliability.

In some studies, qualitative estimates (i.e., ever/never exposed to an agent) or categorical estimates without measurement units (e.g., low, medium, or high exposure level) are developed. Qualitative estimates generally do not require the existence of measurements and fewer resources than more quantitative approaches. Categorical estimates allow the investigation of exposure-response relationships, but are more subject to exposure misclassification than the quantitative approach and, therefore, more likely to miss exposure-disease relationships. On the other hand, quantitative estimation requires measurement data and considerable resources, but it is less susceptible to exposure misclassification and most capable of identifying exposure-disease relationships.

This document describes the background of the spill event and exposure assessment concepts and provides the overall strategy for the GuLF STUDY for characterization and assessment of exposures to constituents of crude oil and other chemical and physical agents (e.g., heat stress and noise) possibly experienced by clean-up workers. It provides information regarding data collection and compilation, identification of the hazards, development of exposure scenarios, data analysis, exposure reconstruction, method evaluation, quality control, data management, and the general time line for completion of exposure assessment activities. The goal of the exposure assessment will be to develop precise (preferring quantitative over qualitative), credible, and well-documented exposure estimates to various chemical and physical agents for each cohort member from three routes of exposure (inhalation, dermal, ingestion). Possible exposure estimates for clean-up activities and for off-hours (i.e., when the subject wasn't working) will use state-of-the-art assessment methodologies.

A large number of air measurements have been collected during clean-up activities by BP and a number of federal agencies. These can be used to characterize workers' breathing zone (i.e., personal measurements), and workplace (area), concentrations for some agents. In additional, a substantial number of offshore and land environmental air measurements have been collected, along with measurements of crude and weathered oil, water, soil, and sediment. At this time, the GuLF Study investigators have not received the original data from these groups and so have evaluated only publicly available information. The original data include additional information that will be of considerable value to the exposure assessment effort. Thus, the final step-by-step approach for exposure assessment cannot be fully specified at this time. It will be developed after review of the available measurement data and assembly of the information on exposure determinants has been completed.

BACKGROUND

Published Data on Other Spills

The Deepwater Horizon oil spill is not the first large spill to occur. One of the first spills to attract worldwide attention was the Exxon Valdez spill in Alaska in 1989 (Gorman et al., 1991). Since that time there have been six other spills that have received the attention of occupational health specialists: the Braer (Scotland, 1993); the Sea Empress (Wales, 1996); the Nakhodka (Japan, 1997); the Prestige (Spain, 2002), the Tasman Spirit (Pakistan, 2003), and the Hebei Spirit (South Korea, 2007) (Table 1). Measurements taken for these disasters have been for THC, VOCs, and BTEX. Levels have all been lower than those measured in oil refineries (Spear et al., 1987) and offshore oil drilling

(Steinsvag et al., 2006; Kirkeleit et al., 2006). Dermal measurements for oil residue were taken for the Exxon Valdez spill. Pre-work shift oil residue levels, however, were higher than the post-work shift levels (Gorman, 1991). In the study on the Prestige spill, metals (aluminum, nickel, lead, and cadmium) in the blood were found at elevated levels compared to controls (Pérez-Cadahía et al., 2008). None of the air or biologic measurements presented in these studies, however, was used in the epidemiologic analyses to evaluate health outcomes, which is the goal for the GuLF Study. In addition, health outcomes were evaluated for a limited number of jobs (e.g., 3 or 4) or for residents located near the spill.

The Gulf Oil Spill

The exposure assessment for the GuLF STUDY can be thought of as comprising five areas: the hot zone (a ~2000 ft radius around the wellhead); the source area, the area within approximately 5 nautical miles of the source; offshore other than at the source; near shore (within site of the shore, i.e., ~3 nautical miles); and land. The land areas includes the beaches, marshes, rock jetties, ports and docks, land-based decontamination areas, and support areas such as eating facilities, offices, training facilities, etc.

In the hot zone, four rig or platform ships repaired the well, collected the oil coming up from the wellhead, and burned or stored it for transfer by barge back to land. Access to the area immediately around the four vessels was highly restricted. There were many other ships and barges in the source area that supported the four rig operation by pumping mud, processing oil, applying dispersant to the wellhead (5000 ft below the surface) and to the water surface, and transferring supplies, personnel, materials, and chemicals to and from ships located in the source area. Offshore operations included three types of vessels: boats, ships and barges. The boats, vessels of opportunity (VoOs), typically were shrimpers, oyster boats and other small boats that laid and collected booms to contain the oil, transported personnel, scouted for oil, supplied personnel, and assisted in skimming and burning operations. Larger boats or ships supplied equipment and fuel to the hot zone and source, skimmed and burned oil on the water surface, carried oil and oily water from the source and hot zone to land, and applied dispersants. Barges carried fuel and raw materials to the source and hot zone and carried oil and oily water to land from the hot zone and source and offshore areas. Barges also cleaned or decontaminated boats of oil offshore. Near shore operations included some of the offshore operations, as well as scouting for oil on the beaches, marshes, and bayous, collecting contaminated wildlife, and cleaning rock jetties and other shoreline structures. Generally, dispersant was not applied near shore or on land. Land activities involved hand and mechanical cleaning of the sand and collection of oil and tar both at beaches and in marshes, wildlife rehabilitation, transport of workers by land vehicles, vessel and equipment decontamination, and support activities, such as material handling, security, provision of food and protective equipment, office workers, and others. Most employees worked for contractors or sub-contractors to BP.

Exposure Concepts

Airborne exposures can be considered from two sources: near field and far field (Nicas, 2009). Near field exposures arise from the performance of a task and thus, the emission is generally within three feet of the individual. Far field exposures arise from sources more than three feet from the individual and can be considered background exposures. Inhalation of some chemicals, e.g., THC, VOCs, PAHs, BTEX, 2-BE, and PG may come from both sources for many GuLF STUDY scenarios. (The term source, here, is the point of emission of the chemical agent of interest, and not the source of crude oil at the wellhead.) For example, workers skimming oil off the water at the source may be receiving airborne benzene exposure from skimming (near field) and from being in the gulf (far field) during skimming, but only far field exposure when not skimming. In some cases, there is only one source. For beach clean-up workers, for example, the inhalation source to BTEX is far field. There is no “near field” exposure to BTEX from the oil because the oil on the beach is weathered and does not produce volatile chemicals, such as BTEX. For each exposure scenario and agent being assessed, a determination will be made as to whether exposure is from near field, far field, or both fields and estimates developed for the relevant fields, if appropriate.

Exposure determinants are factors that affect exposures. They can include factors of the workplace (e.g., indoor vs. outdoors, the presence, and effectiveness, of ventilation systems), factors of the job (e.g., high pressure spraying vs. low pressure spraying during decontamination of boats), and of the individual (e.g., the wearing of PPE). Exposure determinants are important for understanding the differences among measurement data. For example, as part of decontaminating vessels of oil, workers sprayed the vessel hull with detergent using two levels of pressure. The level of pressure can affect exposure levels. The measurements will be reviewed to determine if the two types of sprays used in decontaminating vessels are identified and, if so, if there is a difference between the exposures from the two pressure levels. If there is a difference, two exposure scenarios will be developed: one for high and one for low pressure spraying. If there is no difference and based on other estimation approaches (see Modeling below), no difference is expected; only one exposure scenario would be necessary. This is an example of a determinant that affects a specific job group. Determinants vary by job, which is why understanding of the jobs is crucial to the estimation process. It is unknown at this time how much information is available on exposure determinants. Additional information on exposure determinants will be obtained from the full and complete monitoring datasets when available and from additional discussions with workers and agencies engaged in the clean-up effort.

METHODS

Data Collection

Individual-specific information

The GuLF STUDY questionnaire administered to the cohort members was developed to collect individual-specific information on clean-up jobs (jobs directly involving exposure, such as skimming and beach clean-up, or indirectly involving exposure, such as safety or industrial hygiene support) and areas (e.g., rig/platform ships or ships and boats). Questions on dates, the number of days worked per week, and the hours worked per day, use of PPE, dermal contact with chemical agents, heat stress episodes, and sleeping locations are included in the questionnaire. The questionnaire was developed after reviewing the air monitoring data collected by BP and federal agencies, information contained in numerous documents available on the spill, and interviews of personnel from BP and federal agencies who were involved in the emergency response effort. The information on jobs, including the location and dates worked, will be translated to exposure scenarios (see Development of Exposure Scenarios below), that will be the link between the exposure estimates and the study subjects.

Gulf Oil Spill Historic Exposure Data

There has been a large amount of monitoring information collected on the oil spill. BP and numerous government and non-government agencies have performed measurements, assembled databases, and written reports for the public. Databases identified to date containing measurements of worker airborne exposures are listed in Table 2. Environmental data of air, water, sediment, and weathered oil are also available. NIEHS is currently in the process of developing Memoranda of Agreement with the various government agencies and with BP to obtain access to any additional relevant measurement data, as well as the original sample sheets and field notes.

The GuLF STUDY investigators have met, and will continue to meet, with BP officials (and their contractors) and government representatives to identify other information that can be used to assess clean-up exposures. BP has a number of other potentially useful databases or reports. These include listings of contractors and the type of workers they hired for the clean-up operation, listings of the vessels permitted within a five-mile radius of the wellhead source, clean-up worker sign-in and sign-out sheets at various locations, standard operating procedures, Material Safety Data Sheets for relevant chemicals, action plans (e.g., steps taken when air concentrations exceeded certain levels), and PPE and mobile equipment used.

Another potentially valuable source of information on work activities are reports of worker illnesses and health complaints to various organizations. Those identified so far include US National Institute for Occupational Safety and Health's (NIOSH) Health Hazard Evaluations (NIOSH, 2010a and b), the Louisiana Department of Health and Hospitals, Office of Public Health, the BP Deepwater Horizon Summary of Medic Log Data, and the Venice Branch Infirmary. There may be other state or local organizations that have collected information. The health reports may provide descriptions of jobs or tasks that provide insight into activities that had higher exposures. NIOSH conducted two Health Hazard Evaluations on workers participating in the spill clean-up, in which investigators administered questionnaires to the workers on the activities they performed (NIOSH, 2010 a and b). The detailed information from these questionnaires is being requested. Other non-government agencies, such as universities, have collected information that may be useful to the exposure assessment.

A large amount of information on the spill is available on public websites (Senn, 2010). Although preliminary searches have been done, a systematic approach will be taken to identify sites that contain important information. Data collected are likely primarily to be descriptive information and include pictures, although some measurement data may be found other than those already described.

Current Measurements

Consideration was made for the collection of additional monitoring data. At this time, however, only one clean-up activity is still being performed, that of beach clean-up. We have decided that measurement of this operation is not an efficient use of resources because only 1% of the approximately 9000 beach clean-up measurements from the various organizations on this job were above detectable levels, the frequency of finding tar balls and other oily and tar materials has decreased over time, and the cooler temperatures are likely to result in even lower emissions than in the summer.

Measurements in biologic tissues will be made on the 6000 individuals in the Biomedical Surveillance Cohort (Sandler et al., 2010). Oil contains some persistent chemicals, among which are heavy metals. In the Prestige study blood levels of lead, aluminum, chromium, and nickel were significantly higher in the blood of workers compared with controls (Pérez-Cadahía et al., 2008). Evaluation for such metals is being considered, along with other persistent chemicals, as a component of the biologic measurement effort.

Site Visits and Interviews

Site visits were made in November and December, 2010, to observe vessel decontamination and several types of beach and marsh clean-up activities. Samples of tar balls, tar mats, weathered oil, and oily plant material were collected from the beaches and marshes. Interviews with workers and BP and government employees familiar with various activities, including industrial hygienists, were conducted. A few additional site visits may be made in the future, but emphasis will be on interviewing workers on jobs no longer being conducted, including possible sources of variability.

Compilation of the Data

All documents received will be entered into a document tracking system associated with specific keywords (e.g., oil composition, PPE) to ensure that information is not lost and is easily retrievable.

Measurements

Documentation of text fields in the measurements collected by BP is currently being standardized by BP to ensure that entries on locations, jobs, and other variables are consistent. The original data, however, are not being changed (Hewitt personal communication, 2010). Other databases (such as those from the Occupational Safety and Health (OSHA), NIOSH, and the US Coast Guard) are included in this standardization. Other information not in these databases, such as information on the original field notes or sampling sheets, will be abstracted and entered into the measurement database by the NIEHS industrial hygienists. This information will include measurement descriptors (e.g., temperature and humidity measurements) and exposure determinants, such as location, work practices, and equipment to allow assignment of each measurement to an exposure scenario.

Descriptive Data

The actual measurements provide only a part of the information needed to accurately estimate exposure levels. Descriptive data characterizing the spill will provide insight into working conditions and identify possible variability of these conditions. They can provide information on exposure determinants, such as work tasks, material components (e.g., the fresh oil, weathered oil, the dispersants), distance from the source, quantity of the agent present, application rate, points of emission, frequency and duration of exposure, size and surface area of containers, use of PPE, likely movement of air and water currents based on air and water models, and weather conditions (temperature, humidity, wind speed, and direction). Agent-specific information such as absorption rate (for dermal exposure), vapor hazard index (for inhalation hazards), and PPE permeation rates (i.e., permeation of chemical agents through the PPE) is also important. The descriptive data will be assigned keywords to allow easy retrieval. A database of photographs also will be developed to aid in understanding the jobs and their variability.

It is proposed to use a relational database for the measurement database due to the large number of observations (several hundred thousand) and the need to link these measurement data to other descriptive data, such as job descriptions, locations, operating procedures, PPE, pictures, and environmental measurements. Relational software such as Microsoft Access™ or Oracle™ will be used to maintain these data, with the exposure scenario providing the link to the measurement and descriptive data.

Identification of the Hazards

Over 250 chemical and physical agents have been measured during the spill. Many of these agents, however, were identified from analytic techniques that automatically evaluate a specific array of analytes, rather than being specifically targeted for analysis with the expectation that the agents were present. For many of these agents, reported levels were below the level of detection of the sampling and analytic methods. An example of such an array is the presence of 39 pesticides in the database. These pesticides were not used in the spill effort and there is no reason to assume that most of these were present. Various bug sprays were used, however, and the active ingredients of the sprays are being investigated.

There are, however, a number of chemical agents that were measured specifically because of their presence in oil or because of the jobs being performed. Criteria for selecting which agents will be assessed include the number of individuals exposed; likely exposure levels; the availability of measurement data and other information to assess the exposures; and possible health effects. The list of agents assessed will include VOCs, THCs, BTEX, PAHs, and possibly hydrogen sulfide, from the oil and the weathered oil. The two dispersants applied to the water, 2-BE and PG, and possibly other surfactants used in the dispersant and in the cleaning chemicals, will be estimated. Limonene, one of the major cleaning chemicals used in the decontamination of vessels and equipment, is being considered. Other agents under consideration are: carbon monoxide, gasoline exhaust, and elemental carbon (from engine exhaust); dioxins (from burning oil); phthalates (unknown source at this time); insecticides and sunscreen; metals (particularly nickel, vanadium, aluminum, cadmium, and lead, from the oil); particulate matter (PM2.5 and PM 10 from the oil burning); heat stress; and noise. Evaluation of the physical state of the agents (e.g., aerosol and vapors for the dispersant; fumes and smoke for PAHs) will be made to determine the importance of possible differences in toxicity.

Development of Exposure Scenarios

Typically in epidemiologic cohort studies, such as this study, exposure estimates are developed for jobs and not for individuals, called JEMs. The estimates are then assigned to the study subjects through the job. In this study, exposure groups, i.e., jobs, are being developed based on the air monitoring data and the information collected to date. These groups are job- or task-based, and examples include drillers, moon pool workers (the hole directly over the well on the rig or platform ships) and "other rig workers" on the rig; oil burners and crews on oil burning ships; oily sand clean-up workers, tar clean-up workers, and mobile equipment workers on the beach; and barge and land decontamination workers who in cleaning boats and other equipment used: 1) ice pellets, 2) low pressure spraying of detergents, 3) high pressure spraying of detergents, and/or 4) rags, brushes, and other manual equipment. It is estimated that there are likely to be 50-60 job groups. Discussions with clean-up workers and organizations continue to identify the full list of job groups.

Once job groups have been identified, they will be assessed to identify exposure scenarios. An exposure scenario is a job group with a similar exposure distribution. Different jobs may be grouped because of similar exposures. On the other hand, jobs with the same title may be placed in separate exposure groups because of differing locations or work practices affecting exposure levels. For example, area may be an important determinant of exposure. Skimmers at the source were likely to have had higher exposure levels to BETX because they were closer to the source, and thus the oil was fresher, than skimmers offshore, who more often encountered weathered oil, from which BETX likely had volatilized. Month may have had an effect on exposure levels. Summer beach clean-up workers likely had higher exposure levels than winter clean-up workers. Geographic location (e.g., Louisiana, Florida) also may be important. The oil spill spread out over the gulf and to different points on land at different times. The oil reached the islands near Venice, LA approximately on April 30th, whereas the oil did not reach the Florida beaches until about June 4th. We estimate there will be about a thousand exposure scenarios for which estimates will need to be developed for each agent.

Data Analysis

Monitoring Data Issues

Although there are a large number of measurements, data analysis and exposure characterization will be challenging because of the large number of chemicals, differences in analytical methods and instrumentation employed, limitations of sampling devices, number of non-detectable measurements, use of compliance-based rather than research-based sampling strategies, and severe working conditions. Specific issues are:

1) Various sampling and analytic techniques were used to measure exposures. Many of the sampling methods used to measure personal exposures were developed, calibrated, and validated based on occupational exposure limits, rather than the much lower exposure ranges found in this study. Thus, the levels of many measurements could not be quantified.

2) The (personal) measurement devices worn by individuals were primarily charcoal absorbent sampling badges. Measurements using these devices often were collected under conditions (temperature, humidity, and sample duration) that were outside of the range of manufacturers' recommendations. These conditions can lead to an underestimation of the actual exposure levels.

3) One of the most common detectors used to analyze the personal near field samples collected in the study was the hydrogen flame ionization detector (HFID or FID), which responds differently from the photoionization detector (PID) taken for other study measurements (see point 4). This means that if a personal sample were collected and analyzed using a FID, the results of the analysis would be numerically different from what would be obtained using an instrument with a PID. Many of the environmental measurements were analyzed with a PID instrument, which will be used to determine far field exposure levels. Because exposure levels from both the near and far fields will be estimated and combined into a single value, they must be on an equivalent scale.

4) Other measurements available in the study were collected with direct reading instruments using a photoionization detector [PID]). First, the PID is not sensitive to a number of the agents associated with the clean-up operation, and the response (i.e. relative sensitivity) of the PID varies by agent. Second, in contrast to the FID, which is agent-specific, the PID is calibrated only to isobutene at 100 ppm. This level is well above the observed volatile organic carbon (VOC) concentrations of <10 ppm observed for most measurements outside the hot zone. In addition, many of the agents of interest in the study have response sensitivities different from and well below isobutene response. These two issues (sensitivity and calibration) probably resulted in a considerable underreporting of agent concentrations.

5) The sampling periods of the measurements differed. The large number of non-detectable measurements may be the result of limitations of the sampling methods and the short sample

periods. First, each sampling and analytic method used in the study has its own limit of detection (LOD) and limit of quantification (LOQ) below which the measurement results are not reliable. Second, shorter sampling periods result in higher LOQs and LODs. Thus, the LOD and LOQ are directly related to the sampling and analytic methods and the sampling period, so that measurements on the same activity from different methods and of durations may have different LODs and LOQs.

6) BP used a compliance-based sampling strategy. The goal was to identify the higher exposed jobs. Other job groups were not directly assessed and thus have no measurement data. Assessment of exposure levels for these groups requires measurement data from similar jobs and estimation of the critical exposure determinants.

7) A number of unusual circumstances were experienced by the workers that are not typically encountered when measuring air concentrations. Workers' schedules were typically 12 hours per day, 7 days per week, with few days off over a significant period of time, such as weeks, reducing the body's recovery time. Although these circumstances do not affect the measurements, they do affect the interpretation of the measurements.

All of these issues complicate the evaluation and interpretation of the monitoring results. The approaches for handling these situations will have to be determined prior to development of the exposure estimates.

Preliminary Statistical Analysis

After resolving the sampling and analytic issues just described, preliminary analyses on the measurement data can be completed.

1) Because of the limited number of measurements and the large number of non-detectable measurements, the relationships between the various chemical agents will be evaluated to determine if a marker agent exists. Preliminary analyses suggest that there may be correlations between various agents, such as BTEX and THCs. If so, these correlations can be used to estimate exposures to some agents that lack sufficient (detectable)measurements.

2) Information on the minimum detectable and the minimum quantitative levels of the sampling and analytic methods will be used with statistical techniques to estimate exposure levels below the LOD and the limit of quantification (LOQ).

3) Descriptive statistics, such as means, geometric means, geometric standard deviations, and the 95th percentile of the measurements, will be calculated.

4) The measurement data will be analyzed within each exposure scenario to characterize exposure patterns, identify outliers that may skew exposure metrics, and identify the possibility of more than one exposure distribution (which may suggest that two or more exposure scenarios have been inappropriately grouped).

5) The measurements will be analyzed to identify important exposure determinants. Preliminary analysis on the publicly available BP data, suggest that the area (e.g., hot zone, source, offshore, near shore, and land) appears to be related to the number of detectable measurements. For near shore and land activities, the month of exposure appears to be an important contributor to the variability of the measurements, but this does not seem to be true for offshore activities.

Exposure reconstruction

Determination of the Assessment Method

The goal of the exposure assessment is to develop inhalation, dermal, and ingestion exposure levels of specific chemicals (in crude oil and from other sources) that were encountered by individuals both on and off the job during the clean-up operations. This effort will be challenging due to the large number of non-detectable measurements, the lack of dermal measurements, the need for a large amount of descriptive data, the large number of agents being assessed, and the possibility of exposure through several routes. In addition, although the total number of measurements is large, there may be some exposure scenarios with none or only a few measurements. The approach will be to develop exposure scenario matrices (similar to JEMs) that identify exposure levels of each scenario for the various agents. These estimated levels will be assigned to each individual study subject by way of the subject's exposure scenario history and then modified based on the individual-specific information collected from the questionnaire (such as PPE use).

Estimated exposure levels will be developed for scenarios as precisely as possible given the resources available. Estimation of quantitative levels will be the preferred approach, but if the data are not sufficient for this approach, exposure categories, either with or without measurement units (e.g., mg/m^3), or even qualitative (ever/never) estimates, will be developed. The decision as to which metrics to employ will be based on the number of relevant measurements, the number of detectable measurements, and the ability to apply mathematical (using physical and chemical properties), statistical, and deterministic models (see Modeling of Inhalation Exposure Levels).

In particular, for any given agent, the goal will be to first estimate a quantitative level of exposure for each exposure scenario. If a quantitative assessment is not possible or practical, a categorical assessment will be made. If that is not possible or practical, a qualitative (yes/no) evaluation will be made. As a result, some scenarios may have quantitative estimates for an agent, but others may have semi-quantitative estimates for the same agent, and some may only be categorical. For example, it may be possible to estimate THC exposure in ppm for most exposure scenarios. For workers with these estimates, exposure-response relationships can be investigated using the continuous data. For other scenarios, however, only categorical estimation may be possible. For example, for some scenarios, the exposure-response relationship may be based on categorical ranges of exposure, such as <0. 1 ppm, 0.1-1 ppm, and >1 ppm. Thus, the approach may result in estimates for various scenarios having different levels of precision for the same agent. It will be necessary, therefore, to convert the more quantitative estimates to less quantitative estimates to be able to establish a common level of precision (i.e., a common denominator) across all scenarios. This is necessary to allow investigations of exposure-response relationships among all study subjects regardless of the precision of the assessment.

For some agents, quantitative estimates will not be possible. For noise, for example, there are essentially no measurements, so that the estimation is likely to be categorical (e.g., <90 dbA, 90-100 dbA, and >100 dbA). In the case of insecticides and sunscreen, estimation is likely to be qualitative only, e.g., used/not used, as reported by study subjects in the questionnaire (with frequency of use used for the exposure-response investigation). Estimation of dermal and ingestion hazards are likely to be categorical only, due to the limitations of current dermal and ingestion exposure assessment techniques. (Ingestion hazards only will be estimated from accidental contamination of the mouth or accidental ingestion due to deposition on food from contaminated hands, forearms or face, not from food or drinking water ingestion.)

Modeling of Inhalation Exposure Levels

Various modeling methods will be used to estimate exposures for scenarios lacking sufficient monitoring data. For example, mathematical models with exposure data from one agent can be used to estimate the exposure levels for another agent if the chemical and physical properties of the two agents are similar. Knowing the exposure level for one chemical agent in one task, such as VOCs from high pressure cleaning of vessels, may provide sufficient information to estimate exposure levels to another chemical agent in the same mixture, such as benzene, using vapor pressure. Knowing the exposure level of limonene from high pressure spraying also may allow estimation of exposure levels of limonene from low pressure spraying when the application rates of the spraying are available (Stenzel, 2006).

Determinant models will be used to estimate exposure levels where necessary and possible. These models use measurement data to predict the influence of a determinant on exposure levels that can then be used in other scenarios without measurement data with the same determinant. For example, if analyses of monitoring data finds that distance from the hot zone decreases far field exposure levels in a particular direction by a predictable amount (accounting for meteorological conditions), this information could then be applied to all workers in the far field in all directions.

Use of uncertainty analyses also is being considered because information on determinants is not error free. For example, Material Safety Data Sheets often report ranges of the percentage of various components the product contains, rather than a single point value. Monte Carlo or similar types of simulation techniques will be considered for estimating the uncertainty in material formulations in such instances.

Bayesian statistics is another estimation approach that may provide exposure estimates. Bayesian statistics are useful for small datasets or where there is some prior knowledge of an exposure distribution. For example, if, for a given exposure scenario, sufficient measurement data are available for estimating exposure levels for June, July and September, but not for August, Bayesian statistics may provide an estimate for August. Similarly, the availability of sufficient data in one location for an exposure scenario may provide useful information for the exposure levels of a scenario in a different location.

Both the BP personal samples and the Environmental Protection Agency (EPA) measurements contain latitude and longitude data, which, when combined with airflow dynamic models, may allow estimation of air concentrations at the various locations in the gulf and along the gulf shore over time. Geographic information systems (GIS) or spatial statistics using such data may be helpful in estimating exposure levels for individuals for far field exposures.

Modeling of Dermal and Ingestion Exposures

There was substantial opportunity for dermal exposure during the spill clean-up. Rig/platform workers were likely to have had hand contact with oil and oily water. Burning of oil likely resulted in smoke particulates landing on the skin. Beach workers likely picked up contaminated rakes and other sand cleaning equipment. Although protective equipment was worn by these workers, permeation of the oil and other agents through the equipment, resulting in exposure, needs to be considered. The lack of dermal measurements means that estimation will require the use of modeling.

There is little in the published literature on estimating dermal exposure levels (Vermeulen et al., 2002). The most widely recognized and evaluated model is the DeRmal Exposure Assessment Model (DREAM, van Wendel de Joode, 2003). This model uses exposure determinants in an algorithm to develop relative differences among dermal exposures (called DREAM units). Estimation of dermal exposures for the GuLF STUDY will be done based on the DREAM algorithms that use the physical and chemical properties of the agents and the method of contact (emission, e.g., direct contact with the agent through splashing or handling, such as the rig workers above), deposition, e.g., landing on the skin from the air, such as the oil burning workers above; and transfer, e.g., landing on the skin through coming into contact with a contaminated surface, such as the beach workers above). Other components of the model are the part of the body contaminated (e.g., head, hands, forearms, etc.), frequency of contact location, and the use of protective equipment. The model has been found to distinguish among exposure scenarios with differences in expected levels (de Joode *et al*., 2005). Ranking of the study subjects by DREAM units, although not based on a measurement scale, is useful for epidemiologic analyses. It is also possible that actual ug/cm^2 levels could be estimated by relating the DREAM units developed here to previous DREAM studies or by relating specific exposure scenarios to published dermal monitoring studies (such as relating spraying during vessel decontamination with pesticide spraying).

People typically touch their face regularly. Under the particular circumstances of the spill, i.e., the high temperature and humidity, face touching was likely to be substantial due to wiping the face of sweat. If this is done with hands covered with contaminated gloves, the areas around the mouth are likely to have become contaminated. There are no established methods for estimating ingestion hazards, but some work has been done (Cherrie, personal communication, 2010). Attempts to estimate ingestion levels of substances around the mouth and from eating with contaminated hands are being considered.

Other Issues

Exposure estimates will be developed accounting for the absence or use of PPE. The effect of weathered oil, whereby the more volatile components have evaporated, will be considered. Consideration will be given to estimating exposures based on body burden (accounting for all routes of exposure) and for long work days, which may reduce the body's recovery time. Estimates will also be developed for off the job exposures to the same analytes, since most of the workers slept in areas near their day jobs (e.g., in the hot zone [e.g., the crews of the rigs and supporting vessels], offshore [in flotels or on barges], or on land in nearby communities.)

Exposure Estimation Units

At this time, it is anticipated that exposure scenarios will be developed based on job groups and tasks, area (hot zone, source, offshore, near shore, and land), month of exposure, and geographic location. We used these scenarios as the basis for questions in the questionnaire and, therefore, the responses are likely to be directly linkable to exposure estimates. Information on geographic location and dates performing various jobs also is also being collected in the questionnaire. The large number of scenarios, the number of analytes (perhaps 20), the possible time periods (perhaps 10 months, from April to February of 2011 when the interviewing will start), and perhaps 20 geographic locations for inhalation, dermal and ingestion hazards indicates that a large number of estimates must be developed. These will be combined wherever possible to make the estimation process manageable. Although much of the estimation will be mechanized once one scenario has been estimated, the exposure assessment is a sizeable, complicated, time consuming, and challenging task.

Because estimating exposures relies on differing types of information with differing quality, an estimate of the confidence will be developed to provide some indication of the information's reliability. For example, where estimates from various models result in a similar ranking of a particular exposure scenario for a particular agent, the confidence of the estimates might be rated as high. If different models result in very different rankings for a scenario, the confidence in the estimates might be rated as low. This procedure can allow low confidence individuals to be excluded from the epidemiologic analysis.

Exposure Metrics for the Epidemiologic Analysis

The specific exposure metrics to be developed are listed in Table 3. It may be necessary to determine these exposure metrics separately for full-shift, for short-term and for peak exposures. For example, a worker may have received all of his/her exposure from performing a task that took only one hour per day. A measurement taken for the full twelve-hour shift would result in a reported measurement result of only one-twelfth of the exposure experienced during the one hour task, and thus could result in missing an association to a particular disease. Information on peaks may be derived from the short duration samples or the direct reading instrument measurements.

Evaluation of the Exposure Assessment Methods

Validation of the exposure assessment process is crucial to the credibility of the study. Validation, however, requires a gold standard against which to compare information provided by study subjects or estimates from the exposure assessment process. Measurements are usually considered a gold standard, but measurement data available for the exposure period of interest are usually rare, limited in numbers, and/or used for the assessment process itself. The use of a limited number of measurement data as a gold standard also is problematic because the representativeness of limited numbers to the true exposure is open to question. Thus, validation of occupational and environmental assessments for epidemiologic studies has been rare in the past. When validation is not possible, to provide some credibility to a study, investigators may evaluate reliability of the assessments. Reliability studies use information that is recognized as not being a gold standard but has some level of accuracy allowing it to be an "alloyed" standard to investigate the confidence one can place in the data (Stewart *et al*., 1996; Stewart *et al*., in press).

This study is no different from most studies in the scarcity of data with which to "validate" exposure estimates. Because of the limitations of the methods and instruments used in this study, the high degree of variation in the working conditions (e.g., temperature, humidity) in the workplace, and other issues associated with the measurements (see Data Analysis/Monitoring Data Issues), the assumption that measurements represent a gold standard may not be true in this study. It may, however, be possible to estimate personal exposure levels of the various scenarios using BP personal, (work) area, and environmental measurements and compare these estimates to the personal measurements collected by the various federal agencies (OSHA, NIOSH, EPA, and Department of Interior). We could also hold back a small percentage of the BP data (e.g., 10%) with which to compare to the estimates.

Consideration is also being given to evaluating the reliability of the estimate by comparing estimates developed from different modeling techniques, e.g., statistical modeling, mathematical modeling based on the agents' physical and chemical properties and deterministic models. If the estimates from the differing models results in similar rankings by exposure of exposure scenarios, confidence in the estimation process is increased. If the models result in substantively different rankings, the epidemiologic analyses can use the differing sets of estimates in a sensitivity analysis. In any case, prior to the evaluation, one set of estimates will be identified as the primary set for use in the epidemiologic analyses.

Sensitivity and uncertainty analyses will be considered. These analyses can be used to evaluate the robustness of the estimates when the value of an exposure determinant is estimated, rather than known. Although one estimate of a determinant may be chosen as the best, other values are possible. A sensitivity analysis can use the possible values for the determinant in the estimation of the exposures and compare the results from the different estimates. Similar estimated levels derived from using different assumptions increases the confidence in the estimates.

Errors in recall bias regarding information from questionnaires are always a concern. To ensure the credibility of the reported information, it may be possible to compare the activities, locations, and dates reported by the study subjects in the questionnaires to activities, locations and dates identified in other records, such as training records.

Quality Control

A critical quality review will be performed on the monitoring data from BP and government agencies. The BP data, being the largest database of personal exposures, are an important source of information. First, a review of the measurements will be performed to ensure that duplicate measurements are not present. Second, it will be necessary to examine the measurement data for bias. If they are biased, the bias can be identified and characterized by scenario and adjusted prior to use in the exposure modeling. Third, because the LOD and LOQ are directly related to sample duration, samples of similar duration will be grouped by exposure scenario, and summary statistics (e.g., arithmetic and geometric mean, geometric standard deviation, etc.) between duration groups will be compared to determine if the summary statistics of the grouped measurements are similar. If so, all measurements can be used in the estimation process, increasing the reliability of the estimates. Fourth, the measurements of each agent/scenario combination will be evaluated to identify possible data outliers (data points that do not fit the data distribution) or the presence of more than one distribution within an exposure scenario.

The exposure scenarios developed will be reviewed by a second industrial hygienist to ensure that the exposure scenarios are unique and have been correctly linked to the measurement data. A sample of any descriptive data used in the models (e.g., determinant information) will also be reviewed by a second industrial hygienist.

Although described above as a method evaluation procedure, comparison of the estimates from the various models can also be considered a quality control check. Where rankings of the exposure scenarios are substantially different, a review of the input data and assumptions will be made to ensure that the differences are not due to errors.

Data Management

All confidential information, such as personal identifying information on sampled workers, will be deleted prior to use of these data by the industrial hygienists. Personal identifying information may, however, be important in evaluating between and within worker variability. In such case, unique identifiers not linked to personal identifying information will be made. All reports will present summary statistics only. All raw data will be retained on NIEHS computers.

Time line

Analyses will begin as soon as the original data are obtained from BP and/or the government agencies. As estimates are completed, epidemiologists can begin analyses. All estimates are expected to be completed by January of 2013.

Completion Date	Task
June 1, 2011	Identification of the Hazards
October 1, 2011	Preliminary Statistical Analysis
September 1, 2011	Data Collection
January 1, 2012	Compilation of the Data
February 1, 2012	Development of Exposure Scenarios
March 1, 2012	Determination of the Assessment Methods
June 1, 2012	Development of the Dermal Exposure Estimates
September 1, 2012	Development of the Inhalation Exposure Estimates
November 1, 2012	Exposure Metrics for the Epidemiologic Analysis
December 1, 2012	Evaluation of the Exposure Assessment Methods
January 1, 2013	Final report

Appendix A
Exposure and Exposure Distributions

When discussing exposure for an epidemiologic study, arithmetic means are usually developed as the exposure metric. Often, the mean is interpreted as the exposure level experienced by a study subject. That interpretation, however, is somewhat simplistic. An individual working in a specific exposure scenario will encounter a range, rather than the same specific exposure level, in the day-to-day experience. This range in exposure is referred to as an exposure distribution. Environmental exposures are log normally distributed, rather than following the typical bell-shaped distribution. A lognormal distribution curve is highly skewed to the right on a frequency graph, i.e., the curve rises sharply on the left (denoting a large number of measurements at low levels) and tails off to the right (denoting a small number of measurements at high levels). This lognormal distribution also is highly variable. Typically the upper bound of the distribution usually represented by the 95th percentile is 10 to 100 times greater than the lower bound usually represented by the 5th percentile. This variability in exposures is due to natural day-to-day variation in the workplace, in jobs performed and in the procedures used to perform these jobs. In addition to day-to-day variability, however, there also is uncertainty, which is related to the limitations of the sampling and analytic methods used to measure exposure levels. There are also limitations of the sampling strategies, which usually results in only a very small portion of the actual work shifts measured, although hundreds of thousands of measurements may be collected (as in this study). Estimating long-term exposure levels are affected by these various limitations.

Consider the following example to visualize the concept of variability and uncertainty of an exposure distribution. Intuitively, we think of an exposure as being constant, but in reality, it is much like other phenomena that have a high degree of variability, such as the weather or the water line on a beach. Consider the water line example and assume that someone would like to build a house at the ocean's shore and would like a high degree of certainty regarding a safe location. At any point in time, the water line could be measured, but this line varies minute to minute due to the waves and over the day and month due to tides. The water line also varies with the weather and the season. In addition to the variability of the water line, there also is uncertainty associated with the measuring techniques. Using a 12-inch ruler with a scale of fractions of an inch will provide a more exact estimate of a small distance than using a yard stick with a scale only of inches. Finally, there is uncertainty in the true, overall mean when only a limited number of measurements are collected and this uncertainty rises as the variability of the water line increases. Nevertheless, even though the water line on the shore is highly variable, if sufficient measurements are collected with an adequate degree of precision and accuracy, science can predict the upper bound of the water level and thus the safe distance to build a house.

These highly variable exposure distributions and issues with the uncertainty of measurements or modeling techniques greatly complicate the exposure assessment process. There are techniques and methods available to circumvent the problems associated with exposure variability and uncertainly, but these techniques are complicated and can be time consuming.

Table 1. Summary of Historical Oil Spill Exposure Assessment

Name of Vessel	Date	Location	Description of Measurement Results	Exposure Assessment Comments, including job descriptions	Reference	Oil Type & Amt Released
Exxon Valdez	Mar 1989	Alaska	No air measurements; dermal measurements (pre & post shift) but results inconclusive	None	Gorman et al., 1991; Palinkas et al., 1992; Palinkas et al., 2004	North Slope (Prudhoe Bay) crude: 38,500 tonnes
MV Braer	Jan 1993	Scotland	Air: THC ave=2.2, highest=6.3 ppm; benzene ave=0.04, highest=0.07 ppm	Residents w/in 4.5 km radius of Garth's Ness and present at any time on or after 5 Jan; Exposed group, outside a building; Unexposed group, inside a building	Campbell et al., 1993; Campbell et al., 1994	Gullfaks: 85,000 T
Sea Empress	Feb 1996	Wales	No exposure measurements	Exposed group: nearby residents; control group: unexposed residents of north coast	Lyons et al., 1999	Light crude: 72,000 tonnes; heavy fuel oil, 360 tonnes
Nakhodka	Jan 1997	Japan	Total hydrocarbons: ave of ave=0.14 ppm, highest=1.5 ppm; benzene: ave=1 ppb; toluene: ave=4 ppb; xylene: ave=1 ppb; suspended particulates: ave of ave=0.02 mg/m3, highest=0.09 mg/m3	No tasks listed for the 4 personal samples during clean up activity.	Morita et al., 1999	'C oil': 6,000 tons
Erica	Dec 1999	France	N/A	N/A	N/A	Heavy fuel oil: 19 000 metric tons
Prestige	Nov 2002	Spain	No exposure measurement results.	Volunteers, seamen, bird cleaners, salaried workers working $\leq$ or > 5 days	Carrasco et al., 2006 ; Suarez et al., 2005	Fuel oil, #2 and #6: 77,033 T

Name of Vessel	Date	Location	Description of Measurement Results	Exposure Assessment Comments, including job	Reference	Oil Type & Amt Released
Prestige	Nov 2002	Spain	V: Total VOC: ave=483; BTEX: ave=197; benzene: ave=134. MW=Total VOC: ave=201; BTEX: ave=94; benzene: ave=50. HPW: Total VOC: ave=38; BTEX: ave=20; benzene: ave=3 ug/m3	Volunteers (V), hired manual workers (MW), hired high-pressure cleaner (HPW) workers	Pérez-Cadahía et al., 2006; Pérez-Cadahía et al., 2007	Fuel oil, #2 and #6: 77,033 T
Prestige	Nov 2002	Spain	See above	Fishermen participating in clean-up operations; 3 groups: less affected, moderately affected, and most affected areas	Zock et al., 2007	Fuel oil, #2 and #6: 77,033 T
Prestige	Nov 2002	Spain	No air measurements; heavy metals in blood (Al, Cd, Ni, Pb, Zn)	See above	Pérez-Cadahía et al., 2008	Fuel oil, #2 and #6: 77,033 T
Tasman Spirit	Jul 2003	Pakistan	Air: Residents and workers: 40–170 ppm VOCs	Exposed: clean-up operation workers for at least 8–10 h per day for six days/week; control: clerical staff, shopkeepers and salesmen	Meo et al., 2008	Crude oil: 28,000 tons
Hebei Spirit	Dec 2007	South Korea	No air measurements	Coast guard officers, soldiers, residents and volunteers	Lee et al., 2010	Crude oil: 12,388 tons

Table 2. Summary of Major Measurement Databases*

Source	Worker /Work Area Sample	Dates	Jobs identified ?	Location (e.g, boat)	City, State	S&A method	Number of analytes	Number of measurements	n >LOD	%	Comments
OSHA	Worker	May: 1 day; June-Aug: most days; Sept: thru 9/6	job title, no task info	variable	yes	yes	18	4463	36	0.8	Lab analysis
OSHA	Worker	May: 5/24 on; most of June	job title, no task info	variable	yes	no	12	176	13	7	Direct reading
OSHA	Worker	July 7	job title, no task info	variable	yes	no	Noise	20	0	0	
OSHA	Worker	June 8, 16	job title, no task info	variable	yes	no	Heat	5	0	0	
Coast Guard	Worker	June 14-July 2	Job title, task info	yes	no	?	8	1390	1169	84	
BP	Worker/ Work area	April 27-Oct 18	some tasks, some job title	no	no	no	6	119269	7420	6	Job v area v environment unclear
NIOSH	Work Area	June 4-23; Aug 10	NA	yes	no	yes	100	1738	495	28	Targeted events; HHEs provide more info

Source	Worker /Work Area Sample	Dates	Jobs identified ?	Location (e.g, boat)	City, State	S&A method	Number of analytes	Number of measurements	n >LOD	%	Comments
NIOSH	Worker	June 4-23; Aug 10	tasks more than job title	yes	no	yes	52	841	271	32	Targeted events; HHEs provide more info
EPA	Environ mental air	April 28-Sept 6	no	unclear	Coun ty, state	yes	4	58112	33672	58	
EPA	Environ mental air	April 28-Sept 18	no	unclear	Coun ty, state	no	16	10706	3605	34	
Fish and Wildlife	Worker	June 4-16	Task	NA	No	No	7	92	9	0	
Fish and Wildlife	Area	June 4-16	No	No	No	No	7	25	25	0	

*From publically available websites.

Table 3. Possible metrics, and their derivation, for the epidemiologic analysis

	Qualitative	Categorical (without measurement units)	Categorical (with measurement units)	Quantitative
Cumulative	Hours *days	Score*hours*days	Mid-point * hours*days	Intensity*hours*days
Average	NA	Cumulative/(hours*days)	Cumulative/(hours*days)	Cumulative/(hours*days)
95th Percentile	NA	Highest category	Highest mid-point	95th percentile based on average
Peak	NA	Score	Mid-point	Intensity

References

Campbell DD, Cox J, Crum K, et al. (1993) Initial effects of the grounding of the tanker Braer on health in Shetland. The Shetland Health Study Group. Brit Med J 307:1251-1255.

Campbell DD, Cox D, Crum J et al. (1994) Later effects of grounding of tanker Braer on health in Shetland. Brit Med J 309:773-774.

Carrasco JM, Lope V, Perez-Gomez B, et al. (2006) Association between health information. BMC Pub Hlth 6:1.

De Joode, B.V., J.J. Van Hemmen, T. Meijster, V. Major, L. London, H. Kromhout (2005). Reliability of a Semi-Quantitative Method for Dermal Exposure Assessment (DREAM). J Exp Anal Env Epid 15:111–120.

Gorman RW, Berardinelli SP, Bender TR Exxon Valdez Alaska Oil Spill (1991) HETA 89-200 & 89-273-2111. Cincinnati, OH: National Institute for Occupational Safety and Health.

Hill AB (1965) The Environment and Disease: Association or Causation? Proc Roy Soc Med 58:295-300.

Jayjock M, Stenzel M, Anthony TR, Arnold SF (2009). Why Model? In Keil CB, Simmons CE, Anthony TR (eds): Mathematical Models for Estimating Occupational Exposure to Chemicals. 2nd ed. Fairfax, VA: American Industrial Hygiene Association.

Kirkeleit J, Riise T, Bratveit et al. (2006) Benzene exposure on a crude oil production vessel. Ann Occup Hyg 50:123-129.

Lee J, Kim M, Ha M, and Chung BC (2010) Urinary metabolic profiling of volatile organic compounds in acute exposed volunteers after an oil spill in Republic of Korea. Biomed Chromatogr 24:562-8.

Lyons RA, Temple RF, Evans D et al. (1999) Acute health effects of the Sea Empress oil spill. J Epidemiol Comm Hlth 53:306–310.

Meo SA, Al-Drees AM, Meo IM et al. (2008) Lung function in subjects exposed to crude oil spill into sea water. Mar Pollut Bull 56:88-94.

Morita A, Kusaka Y, Deguchi Y, et al. (1999) Acute health problems among the people engaged in the cleanup of the Nadhodka oil spill. Environ Res Section A 81:185-194.

Nicas, M (2009) The Near Field/Far Field (Two-Box) Model with a Constant Contaminant Emission Rate. In: Keil, C.B., Simmons, CE, and Anthony, T.R.A (eds) Mathematical Models for Estimating Occupational Exposure to Chemicals, 2nd ed. Fairfax, VA: American Industrial Hygiene Association.

NIOSH (2010a) Health Hazard Evaluation of Deepwater Horizon Response Workers HETA 20100115: Interim Report #1A Evaluation of May 26, 2010, Hospitalization of Seven Fishermen; Interim Report #1B Evaluation of June 45, 2010 M/V International Peace/MV Warrior Dispersant Mission; Interim Report #2 Evaluation of June 8–10, 2010 In-situ Oil Burns; Interim Report #2A Venice, Louisiana Branch Infirmary Log Review; Interim Report #3B Evaluation of June 21-22, 2010 Dispersant Releases from the M/V International Peace; Interim Report # 4B Evaluation of Health Effects in Workers Performing Oil Skimming from Floating City #1, Louisiana, June 19–23, 2010;

Interim Report #6A Evaluation of Acute Health Effects among Attendees at a United States Coast Guard Safety and Administrative Meeting; June 18, 2010; Interim Report #6B Evaluation of response workers hospitalized in Louisiana from May 28 through June 22, 2010. Cincinnati, OH: NIOSH.

NIOSH (2010b) Health Hazard Evaluation of Deepwater Horizon Response Workers HETA 2010-0129: Interim Report 5 Evaluation of Wildlife Cleaning and Rehabilitation Workers; Alabama, Florida, Louisiana, and Mississippi, June and July 2010; Interim Report #7. Evaluation of Shore Cleaning Workers; Alabama, Florida, Louisiana, and Mississippi, July, 2010. Cincinnati, OH: NIOSH.

Palinkas LA, Russell J, Downs MA et al. (1992) Ethnic differences in stress, coping, and depressive symptoms after the Exxon Valdez oil spill. J Nerv Men Dis 180:287-295.

Perez-Cadahia B, Laffon B, Pasaro E et al. (2006) Genetic damage induced by accidental environmental pollutants. Sci World J 6:1221-1237.

Pérez-Cadahía B, Mendez J, Pasaro J, et al. (2008) Biomonitoring of human exposure to Prestige oil: Effects on DNA and endocrine parameters. Environ Health Insights 2:83-92.

Pérez-Cadahía B, Lafuente A, Cabaleiro T, et al. (2007) Initial study on the effects of *Prestige* oil on human health. Environ Intern 33:176-185.

Sandler D, Kwok R, Engel LS, et al. (2010) GuLF Study: Gulf Long-Term Follow-Up Study for Oil Spill Clean-Up Workers and Volunteers. National Institute of Environmental Health Sciences.

Senn E: Tool to Establish Retrospective Worker Exposures for NIEHS Gulf Study.

Spear RC, Selvin S, Schulman et al. (1987) Benzene exposure in the petroleum refining industry. Appl Ind Hyg 2:155-163.

Steinsvag K, Bratveit M, Moen BE (2006) Exposure to oil mist and oil vapour during offshore drilling in Norway, 1979-2004. Ann Occup Hyg 50:109-122.

Stenzel, MR (2006) Mixtures and Non-ambient Conditions. In: Bullock, W.H. and Ignacio, J.S.: (eds): A Strategy for Assessing and Managing Occupational Exposures, 3rd ed. Fairfax, VA: American Industrial Hygiene Association.

Stewart PA, Zey JN, Hornung R, Herrick RF, Dosemeci M, Zaebst D, Pottern LM (1996). Exposure assessment for a study of workers exposed to acrylonitrile. III. Evaluation of exposure assessment methods. Appl Occup Environ Hyg 11:1312-21.

Stewart PA, Vermeulen R, Coble JB, Blair A, Schleiff P, Lubin JH, Attfield M, Silverman DT (in press) The Diesel Exhaust in Miners Study: V. Evaluation of the Exposure Assessment Methods. Ann Occup Hyg.

Suarez B, Lope V, Perez-Gomez B, et al. (2005) Acute health problems among subjects involved in the cleanup operation following the Prestige oil spill in Asturias and Cantabria (Spain). Environ Res 99: 413-424.

van Wendel de Joode B, Brower DH, Vermeulen R, et al. (2003) DREAM: A method for semi-quantitative dermal exposure assessment. Ann Occup Hyg 47:71-87.

Vermeulen R, Stewart P, Kromhout H (2002). Dermal exposure assessment in occupational epidemiologic research. Scand J Work Environ Health 28:371-85.

Viet SM, Stenzel MR, Rennix CP, Armstrong, TW, Couch JR (2008) Guideline on Occupational Exposure Reconstruction. Fairfax, VA: American Industrial Hygiene Association.

Zock J-P, Rodriguez-Trigo G, Pozo-Rodriguez F, et al. (2007) Prolonged respiratory symptoms in clean-up workers of the *Prestige* oil spill. Am J Respir Crit Care Med 176:610-616.