Estimation of Dermal Exposure to Oil Spill Response and Clean-up Workers after the *Deepwater Horizon* Disaster

Patricia A. Stewart[1,*], Melanie Gorman Ng[2], John W. Cherrie[3], Anna Jones[4], Richard K. Kwok[5,6], Aaron Blair[7], Lawrence S. Engel[5,8], Dale P. Sandler[5], and Mark R Stenzel[9]

[1]Stewart Exposure Assessments, LLC, 6045 N. 27th. St., Arlington, VA 22207, USA; [2]School of Population and Public Health, Faculty of Medicine, 3rd Floor, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada; [3]Insititue of Occupational Medicine, Research Avenue North, Riccarton, Edinburgh EH14 4AP, UK; [4]Public Health Sciences, Social & Scientific Systems, Inc., a DLH Holdings Company, 4505 Emperor Blvd, Suite 400, Durham, NC 27703, USA; [5]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, 111 T.W. Alexander Drive – MD A3-05, Research Triangle Park, NC 27709, USA; [6]Office of the Director, National Institute of Environmental Health Sciences, 9000 Rockville Pike, Bethesda, MD 20892, USA; [7]National Cancer Institute, Bethesda, MD 20892, USA; [8]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599, USA; [9]Exposure Assessment Applications, LLC, 6045 N. 27th. St., Arlington, VA 22207, USA

*Author to whom correspondence should be addressed. Tel: +1-703-534-2956; e-mail: stenzelmr@aol.com

Submitted 22 June 2021; revised 28 July 2021; editorial decision 29 July 2021; revised version accepted 23 August 2021.

Abstract

The GuLF STUDY is investigating health outcomes associated with oil spill-related chemical exposures among workers involved in the spill response and clean-up following the *Deepwater Horizon* disaster. Due to the lack of dermal exposure measurements, we estimated dermal exposures using a deterministic model, which we customized from a previously published model. Workers provided information on the frequency of contact with oil, tar, chemical dispersants applied to the oil spill and sea water, as well as the use of protective equipment, by job/activity/task. Professional judgment by industrial hygienists served as a source of information for other model variables. The model estimated dermal exposures to total hydrocarbons (THC), benzene, ethylbenzene, toluene, xylene, *n*-hexane (BTEX-H), polycyclic aromatic hydrocarbons (PAHs), and dispersants in GuLF DREAM units (GDUs). Arithmetic means (AMs) of THC exposure estimates across study participants ranged from <0.02 to 5.50 GDUs for oil and <0.02 to 142.14 GDUs for tar. Statistical differences in the estimates were observed among the AMs of the estimates for some broad groups of worker activities over time and for some time periods across the broad groups of activities. N-Hexane had ranges similar to THC for oil exposures (e.g. AMs up to 2.22 GDUs) but not for tar (up to 5.56 GDUs). Benzene, ethylbenzene, toluene, and xylene, in contrast, were characterized by higher exposure levels than THC for oil (AMs up to 12.77, 12.17, 17.45, and 36.77 GDUs, respectively) but lower levels than THC to tar (AMs up to 3.69, 11.65, 42.37, and 88.18 GDUs, respectively). For PAHs, the AMs were as high as 219.31 and 587.98 for oil and tar, respectively. Correlations of these seven substances to each other were high (>0.9) for most of the substances in oil but were lower for some of the substances in tar. These data were linked to the study participants to allow investigation of adverse health effects that may be related to dermal exposures.

Keywords: *Deepwater Horizon*; dermal exposure, exposure assessment; oil spill; total hydrocarbons

Introduction

On 20 April 2010, the *Deepwater Horizon* oil rig exploded in the Gulf of Mexico, causing almost 5 million barrels of oil to be released into the Gulf waters over the following 3 months. Over 55 000 workers were rostered by NIOSH as having participated in the response and clean-up (NIOSH, 2011). Workers had inhalation and dermal exposures to multiple oil-related compounds, as well as possible exposure to chemical dispersants, $PM_{2.5}$, and cleaning products. Although more than 160 000 air measurements were available to characterize inhalation exposures, no dermal or surface wipe measurements had been collected. Furthermore, few measurements of dermal exposure were available from other spills for exposure characterization.

The Gulf Long-term Follow-up Study (GuLF STUDY), initiated by the National Institute of Environmental Health Sciences (NIEHS), is investigating potential adverse health effects associated with the oil spill response and clean-up (OSRC) (Kwok *et al.*, 2017). As part of the exposure assessment effort, we updated and enhanced a previously published dermal exposure deterministic model (Van Wendel de Joode *et al.*, 2003) to better reflect the contribution of various exposure determinants relevant to the GuLF STUDY (Gorman Ng *et al.*, 2021). Using information from both the study participants and the study industrial hygienists, we estimated exposure to total hydrocarbons (THC), benzene, toluene, ethylbenzene, xylene, *n*-hexane (BTEX-H), polycyclic aromatic hydrocarbons (PAHs) (as a single substance), and (total) dispersants due to these substances' inhalation toxicity and their ability to be absorbed into the skin or adversely affect the skin. (Dispersant refers to chemicals sprayed onto an oil slick on the water surface or injected into the water to break down the oil into small droplets that more readily mix with the water. It is the dispersants' components that are associated with possible toxicity.) This paper describes the methods and the results for the dermal assessments.

An overview of the exposure assessment effort for the STUDY is presented in Stewart *et al.* (2021). Development of exposure groups (EGs) is described in Stenzel, Arnold *et al.*, 2021. The assessment of airborne exposures to THC and BTEX-H is described in Huynh *et al.*, 2021a,b,c; Ramachandran *et al.*, 2021; Groth, Banerjee *et al.*, 2021; and Groth, Huynh *et al.*, 2021. Assessment of other airborne exposures is also reported ($PM_{2.5}$ (Pratt *et al.*, 2021); dispersant aerosols (Arnold *et al.*, 2021) and vapors (Stenzel, Groth *et al.*, 2021); and oil mists (Stewart *et al.*, 2021)).

Background

The *Deepwater Horizon* (*DWH*) oil spill led to a massive effort to contain the spill and clean the Gulf of Mexico waters and shoreline. Most of the OSRC activities were suspected as having resulted in dermal exposure to oil, oily salt water, and tar. Two rigs (the *Discoverer Enterprise* (*Enterprise*) and the *Helix Q4000 (Q4000)*), were involved in mitigating the release, capturing the leaking oil/natural gas mixture; and separating the gas from the oil and flaring the gas (*Enterprise*) or the oil/gas mixture (*Q4000*) Huynh *et al.*, 2021a. These rigs were located within 1 nautical mile (nmi, 1.8 km; 1.1 mi) of the wellhead, approximately 50 nmi (93 km) southeast of the Louisiana (LA) shore. Two other drilling rigs, the *Development Driller II (DDII)* and the *Development Driller III (DDIII)*, located within the 1 nmi radius of the wellhead (referred to as the hot zone), were each responsible for drilling a relief well.

Supporting these four vessels was a sizable, but unknown, number of other large marine vessels (MVs) (Huynh *et al.*, 2021c; Ramachandran *et al.*, 2021). Fourteen MVs that piloted remotely operated vehicles

(ROVs), called here ROV vessels, performed several underwater activities, such as moving equipment, collecting water samples and taking videos. Other MVs provided other types of support, for example, pumping fluids into the well for well closure attempts; spraying water onto the flaring vessels to reduce temperatures; storing and transporting collected oil; supplying materials/chemicals/crew; and spraying dispersant onto the water's surface near the rigs. We called the 5 nmi (9 km) radius around the wellhead, excluding the hot zone, where most of the supporting vessels worked, the source.

Research vessels collected samples of water and oil, took water quality measurements, monitored the oil plume and collected oiled and dead wildlife. Other, typically smaller, vessels, such as fishing and shrimping boats, scouted for oil; deployed/maintained/retrieved booms; skimmed or burned oil off the water surface; collected wildlife; carried personnel, equipment and supplies to and from the wellhead area; and transported collected oil and oily water back to shore Huynh *et al.*, 2021b. Vessels also decontaminated (deconned) other vessels, jetties and other manmade structures using pressure spraying. Operations occurred throughout the Gulf of Mexico, north of the wellhead off the LA, Mississippi (MS), Alabama (AL), and Florida (FL) coastlines.

Activities on land (Huynh *et al.*, 2021c) included loading/unloading of everything related to the vessels' functions described above at ports and docks, primarily in LA, MS, AL, and FL. Deconning of vessels, equipment on the vessels, boom and other gear was done by low- or high-pressure spraying or by hand (e.g. rags or absorbents). Beaches and marshes were systematically patrolled to evaluate locations needing attention, after which they were cleaned by picking up oil; oily water; tar balls, patties, and mousse; oiled plants; and garbage by hand or using handheld equipment for disposal in bags. Wildlife was captured and rehabilitated. Support staff included material handling, cooks, housekeeping, office, and security workers.

As the oil was released from the well, it changed composition, due to natural processes including evaporation, dispersion, emulsification, dissolution, photo-oxidation, sedimentation, and biodegradation. These processes occurred from the time the oil was released in the subsea water and continued while it was on the water surface and on land. This change in oil, called weathering, resulted over time in a differential decrease of the percentages of the volatile chemicals and a differential increase in the percentages of the semi- and non-volatile chemicals in the oils and tars (Stenzel, Arnold *et al.*, 2021).

Methods

After review of several dermal assessment models, the DREAM model (Van Wendel de Joode B *et al.*, 2003) was selected as the most appropriate model for the GuLF STUDY because it allowed the use of information on determinants of dermal exposure, some of which had been collected from study participants through a telephone interview. Validation work conducted by the developers of the original model found that the dermal exposure units (DREAM units, a dimensionless unit) correlated well with hand exposure measurements across a range of work sites and exposure agents, but less well with other body parts (Van Wendel de Joode B *et al.*, 2005).

We updated the DREAM model to reflect more recent studies and adapted it, primarily by changing weights of the various determinants to fit the oil-related chemicals and dispersants and available data of the GuLF STUDY to allow retrospective evaluation using interview data from study participants (Gorman Ng *et al.*, 2021). We called this model GuLF DREAM and the dimensionless outcome measures, GuLF DREAM units (GDUs). The model required variables for:

- the concentration, vapor pressure score, and viscosity score of each substance;
- each of three exposure pathways: emission (direct contact with the liquid substance), deposition (contact with the airborne substance), and surface transfer (contact with a contaminated surface);
- the surface area of each of nine body parts (head, upper arm, lower arm, hands, front of torso, back of torso, upper legs, lower legs, and feet);
- for each body part, the intensity (i.e. the percentage of surface area covered) and the frequency of exposure for the likeliest pathway for oil and tar;
- water (seawater removes water soluble compounds) contact; and
- the type and frequency of replacement of gloves and protective clothing.

Each variable was assigned a value, generally between 0 and 5 (Gorman Ng *et al.*, 2021).

To evaluate our model, exposures measured in two studies of heavy fuel oil (in a variety of industries unrelated to oil spills (Christopher *et al.*, 2007)) and of asphalt (among paving workers (Cavallari *et al.*, 2012)) were estimated using GuLF DREAM and compared to those studies' measurements. A correlation (ρ) of 0.59 was found between the GuLF DREAM estimates and the measurements for the hands (Gorman Ng *et al.*, 2021).

There were insufficient measurements to evaluate other body parts.

Because dermal exposure was likely to have occurred both to oil and to tar, we assessed a variety of chemicals contained in each of those two substances, i.e. THC, as petroleum hydrocarbons; each of the BTEX-H chemicals; and PAHs as a single substance. For dispersants, based on the relative quantity of each dispersant used, we evaluated a 33.4/66.6 mixture of COREXIT™ 9527A and 9500A (based on the amounts of the two dispersants used) applied by air across the Gulf and 9500A alone at the wellhead, either sprayed on the water or injected near the seabed.

Data collection

In the GuLF STUDY telephone interview, participants (*N* = 24 937) were asked about the variety of jobs/activities/tasks presented in **Background**, as well as questions that addressed dermal exposure specific to most of the reported job/activities/tasks (for exceptions, see below). A subset of the participants (*N* = 11 193) later completed a home visit interview that included other dermal questions. The dermal questions asked were:

- for each activity reported by the participant, if his/her skin or clothing had contact with each of several substances, including oil, tar, dispersants, and water. (We were unable to find a definition for the difference between oil and tar that we thought the participants would be able to distinguish. We therefore defined the two substances as "a solid or gooey oily residue or tar" (asked first) and "oil or oily water").

If contact occurred, additional questions were asked for each reported job/activity/task:

- the frequency of the contact (none; less than half the time; about half; more than half; or all the time);
- whether the substance got on the hands (yes/no);
- the number of hours a day it was on the hands before being washed off (h, min);
- whether the substances got on the skin or clothing other than the hands (yes/no);
- frequency of contact with water (none; less than half the time; about half; more than half; or all the time); and
- the frequency that each of the eight other body parts became wet with a (unspecified) chemical (<1 day per month; 1–4 days per month; 1–5 days per week; or almost every day), whether on clothing or on the skin.

Questions on the participant's use of protective equipment by job/activity/task covered the use of:

- leather, cotton, or synthetic gloves; and if a second glove was worn;
- boots or rubber slip-ons;
- protective coveralls such as Tyvek;
- long sleeved shirts, jackets, or coveralls; and
- the frequency of changing each of these types of clothing.

All questions were asked of all study participants, except for the set of questions on the frequency of exposure to an unspecified chemical for the eight body parts, which was only asked of the home interview participants. See the online Supplemental Material (SM), Table S1 for more details.

Processing of information for the estimates

Data coding

A set of rules was developed to translate the question responses into weights for the GuLF DREAM model (Gorman Ng *et al.*, 2021). The same weights were assigned for oil, tar and dispersants. For example, if a participant said yes to "Did your skin or clothing come in contact with (substance) during any of your oil spill clean-up work?", a value of 1 was assigned if the response was yes and 0 if the response was no. If yes, a follow-up question was asked, "On an average workday, how much of the time was your skin or clothing in contact with (substance)?" Responses were coded on a scale of 1–5: "None" = 1, "<1/2" = 2, "About ½" = 3, ">1/2" = 4, and "All of it" = 5. Rules were developed based on the substance being assessed, the question number, and the variable.

Questions were missing responses for approximately 5–10% of participants. In addition, only approximately one-third of the cohort, i.e. those who had home visits, had responses on frequency of contact to the eight body parts. To impute information from both types of missing data, a job-exposure matrix (JEM) approach was taken. The basis for the JEM was the set of EGs, unique combinations of jobs/activities/tasks, location, and time (Stenzel, Arnold *et al.*, 2021), developed for inhalation exposures. Time periods (TPs) are described in the SM, Table S2 and jobs/activities/tasks in the SM, Table S3. Briefly, the events in each time period were:

- TP1a (22 April through 14 May 2010): oil flowed from the damaged well. Drilling started on a relief well. Water clean-up activities started. Oil reached the LA shoreline, and beach cleanup started. Dispersant application began.

- TP1b (15 May through 15 July 2010): oil flow continued. Drilling began on a second relief well. Dispersant operations continued. Water clean-up activities continued. The well was successfully mechanically capped on July 15, which essentially stopped the release of oil. Beach and wildlife clean-up was being done in all four states.
- TP2 (16 July through 10 August 2010): the well was "static killed" on 10 August. Water activities started diminishing. Beach and jetty clean-up and decontamination of vessels and equipment continued.
- TP3 (11 August through 30 September 2010): large-scale final decontamination of the vessels and related equipment started. Water activities continued to lessen. Beach clean-up continued but started to decline.
- TP4 (1 October through 31 December 2010): water efforts essentially were completed by the end of December. Beach and marsh clean-up continued by decreasing numbers of people.
- TP5 (1 January to 31 March 2011) and TP6 (1 April to 30 June 2011): beach clean-up continued by decreasing numbers of people. Water operations were limited to the near shore transfer of equipment, supplies, collected materials and personnel. TP6 is distinguished from TP5 because of the warmer ambient air temperatures.

Most of the inhalation EGs were retained, although we modified some to incorporate additional information obtained from the interviews by adding groups to account for differences relevant to dermal exposures. For example, we had an inhalation EG for "Handled/cleaned wildlife", which was as precise as the air measurement data allowed. For the dermal assessment, however, we distinguished among the various wildlife handling/cleaning activities, i.e. "Handled oily wildlife", "Cleaned wildlife", "Used soaps to clean wildlife" and "Retrieved dead wildlife". New EGs were also developed for workers on the rig vessels, because the job-based inhalation EGs had small numbers of respondents. Small sample sizes would have resulted in less precise values when using study participants' responses to impute missing data, so we combined the various jobs to form 7 broad "jobs". Descriptions of all jobs/activities/tasks are provided in SM, Table S3.

The data used to complete the JEM cells were summaries of the responses to each dermal-related question for each EG. Percentages were used for imputing yes/no questions (e.g. if ≥50% of participants answered "yes" to "Did your skin or clothing come in contact with oil during any of your oil spill clean-up work?", the assigned JEM value was 1 (yes), and 0 (no) if <50% of participants answered "yes"). Median integer scores were used for categorical or continuous responses (e.g. for the question, "On an average workday, how much of the time was your skin or clothing in contact with oil?", the responses were coded as: "None" = 1, "<1/2" = 2, "About ½" = 3, ">1/2" = 4, and "All of it" = 5. The integer of the median score across all respondents was assigned as the JEM value.

To minimize participant burden, questions were not asked about changes over time. The JEM values for a given job/activity/task/location, however, varied based on the responses of the participants who worked in each time period. For example, workers who reported cleaning oil pools in LA in TP1a ($N = 59$) had a median emission frequency of 3 ("1/2 the time"), whereas for those who cleaned oil pools in TP1b ($N = 461$), the median emission frequency was 1 ("never"), although many of the 59 in TP1a also worked in TP1b. When <5 participants or <10% of the participants provided responses to an EG, the reported summary value was overridden and replaced by the summary value for "All states" for that activity and time period.

Some GuLF DREAM variables were not sought from the study participants because the variables required information the participants were unlikely to know. Instead, the study industrial hygienists supplied the information based on the scientific literature, knowledge of exposures, extensive published and unpublished documentation on the *DWH* event, and consensus of the study hygienists. This information included three properties of the substances distinguished by the degree of weathering: i.e. the concentration of the various components in the oil or tar; a vapor pressure score; and a viscosity score. The values for these scores and their derivation are described in SM, Table S4. We also identified the percentage of oil weathering associated with each EG (SM, Table S5).

We assigned other exposure determinants. First, we assigned the primary pathway of exposure (emission or surface transfer) based on knowledge and photographs of the job/activity/task. We considered that the contribution of deposition was very small in light of the airborne concentrations measured relative to the contribution from emission or surface transfer (the average of the THC air vapor concentrations across EGs (Huynh *et al.*, 2021a,b,c); Ramachandran *et al.*, 2021 was 0.8 ppm). Thus, deposition was not considered a pathway and deposition frequency and intensity were assigned "0".

Second, the study industrial hygienists entered values for the intensity received on each body part for the appropriate pathway, i.e. emission or surface transfer.

The DREAM definition of intensity was the amount of the body part exposed, and we retained that definition for GuLF DREAM. Intensity values and weights were: <10% of body part = 1; 10–50% of body part = 3; ≥50% of body part = 10. A single intensity weight across all time periods was assigned to each body part for each job/activity/task/location.

Third, some types of clothing and wind speed values were assigned by the industrial hygienists. We assumed headgear was cotton hats. We assumed that workers in all jobs/activities/tasks had been required to wear heavy (leather, cotton, or synthetic) gloves, except for security, general environment/land, housekeeping, kitchen, and office workers, all of whom were assumed to have worn light (latex) gloves. Frequency of hat replacement was considered less than daily and clothes and shoes (rubber booties), daily. The frequency of glove replacement, however, was provided by the study participant: "less than daily", "daily" or "within a work shift". Wind speed was coded as 1 (i.e. no effect on the substance), except for dispersant, which was assigned 0.75 (small effect).

Finally, participants were not asked about dermal exposures for jobs/activities/tasks that were expected to have no or very little dermal exposure, such as "Cooks", "Office workers", and "Security". Other workers were linked to an EG from a response to an open-ended question ("What else did you do?") and therefore did not get asked dermal questions (e.g. workers on ROV and research vessels). The study industrial hygienists entered a single value for each variable across all states and time periods to allow prediction for these jobs/activities/tasks.

Once all participants had complete information for each job/activity/task, either from the responses or the imputed data, we applied the GuLF DREAM model to develop exposure estimates for each participant's unique combination of job/activity/task, location and time period. Because multiple jobs/activities/tasks were reported (median=6 per participant), multiple estimates were assigned per time period.

Statistical analyses

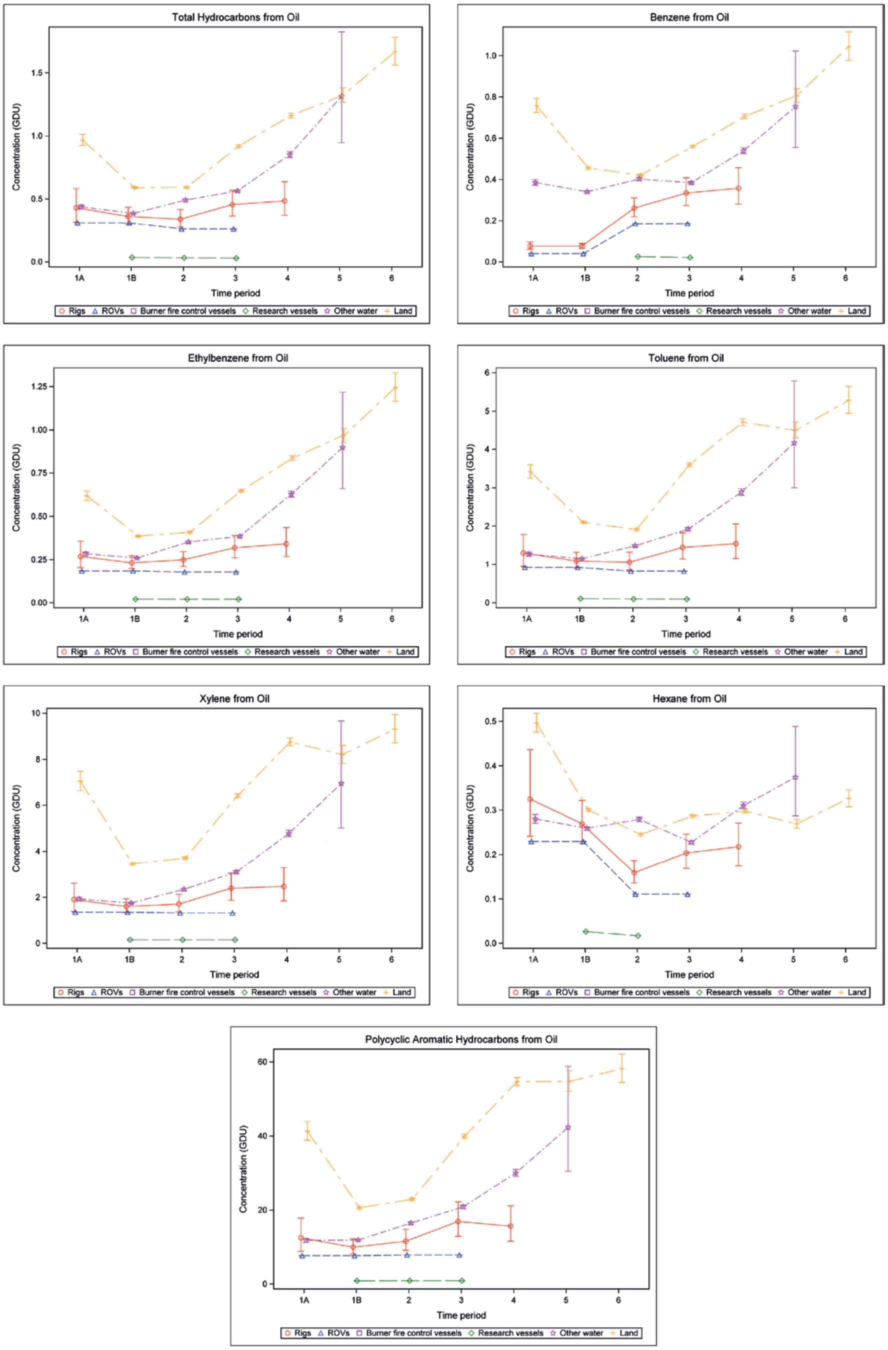
For presentation purposes, the estimated arithmetic means (AMs) of the GDUs across all study participants were calculated by job/activity/task, location, time period and substance. In the AM calculation, we dropped all job/activity/task, location, time period combinations that were <0.01 GDUs. For the graphs, we used the broad groupings of workers for "All rigs", "All ROVs", "All research vessels", "Burner fire control vessels", "Other water", and "Land" by time period for "All states" for oil and for tar. Due to our inability to distinguish between oil and tar under weathering conditions of 25–30%, we combined the oil and tar estimates for any participant's job/activity/task, location and time period associated with those degrees of weathering. This procedure primarily affected the latter time periods (TP3–6). Statistical differences in the AMs were identified by non-overlapping 95% confidence intervals (upper confidence level, UCL; lower confidence level, LCL). Pearson correlations were calculated among the various components of oil and of tar.

Results

The range of THC dermal AM estimates from oil exposure was broad, ranging from AMs <0.02 GDUs (among several groups, for example, "IH/safety-water", "All states", TP1a) to 5.50 GDUs ("Deconned booms/land", "LA", TP3) (not shown). Equivalent tar AMs were <0.02 GDUs (such as "Ran mechanical equipment/ports & docks" All states, TP3) to 142.14 GDUs ("Retrieving boom in shallow water", MS, TP6).

For THC, no statistical differences in oil exposures occurred across time periods among the workers on the rigs (Fig. 1 (note that the scales differ by substance) and SM, Table S6) The values for the variables associated with the ROV, burner fire control and research vessels were assigned by the industrial hygienists and therefore there was little variability in the estimates. Other water workers were characterized by increasingly higher (and statistically significant) mean dermal estimates over time: AM_{TP1b} = 0.39 GDU (LCL: 0.38, UCL: 0.39 GDU); AM_{TP2} = 0.49 GDU (0.48, 0.50 GDU); AM_{TP3} = 0.56 GDU (0.56, 0.57 GDU); AM_{TP4} = 0.85 GDU (0.83, 0.88 GDU); and AM_{TP5} = 1.31 GDU (0.95, 1.83 GDU). For land workers, the AMs significantly fell from TP_{1a} (AM = 0.97 GDU (0.92, 1.01 GDU)) to TP_{1b} and TP_2 (AM for both time periods = 1.59 GDU (0.58, 0.60 GDU)) and then started rising, with statistically significant higher AMs in each of the later time periods (AM_{TP3} = 0.92 GDU (0.91, 0.93 GDU); AM_{TP4} = 1.16 GDU (1.14, 1.18 GDU); AM_{TP5} = 1.32 GDU (1.27, 1.38 GDU); and AM_{TP6} = 1.67 GDU (1.56, 1.78 GDU).

Tar exposure was assessed only for "Other water operations" and "Land", as the other vessels (rigs, ROVs, burner fire control vessels and RVs) had left the area by the time the oil had likely weathered to tar. Water workers were exposed to higher statistically significant differences over time (AM_{TP3} = 0.37 GDU (0.35, 0.38 GDU); AM_{TP4} = 0.86 GDU (0.83, 0.89 GDU); AM_{TP5} = 1.13 GDU (1.05, 1.21 GDU); and AM_{TP6} = 1.46 GDU (1.31, 1.62 GDU)) (Fig. 2 and SM, Table S6). Land workers' tar exposures fell from TP1a to TP3



Total Hydrocarbons from Oil

Benzene from Oil

Ethylbenzene from Oil

Toluene from Oil

Xylene from Oil

Hexane from Oil

Polycyclic Aromatic Hydrocarbons from Oil

Figure 1. Modeled dermal exposure estimates of oil components (total hydrocarbons, benzene, ethylbenzene, toluene, xylene, *n*-hexane, and polycyclic aromatic hydrocarbons by broad activity groups by time period). Note that the scales (in GDUs) vary among the substances. ROVs = vessels piloting remotely operated vehicles. GDU = GuLF DREAM unit.

(AM_{TP1a} = 2.34 GDU (2.00, 2.73 GDU); AM_{TP1b} = 1.28 GDU (1.25, 1.31 GDU); AM_{TP2} = 1.09 GDU (1.06, 1.12 GDU); AM_{TP3} = 1.09 GDU (1.08, 1.11 GDU)), but exposures increased in TP4 and rose substantially in the last 2 time periods (AM_{TP4} = 1.17 GDU (1.15, 1.19 GDU); AM_{TP5} = 11.64 GDU (11.27, 12.02 GDU); and AM_{TP6} = 14.37 GDU (13.68, 15.09 GDU)).

In a comparison of the AMs across the broad groups, land workers generally had statistically higher exposures than did other water workers, who generally had

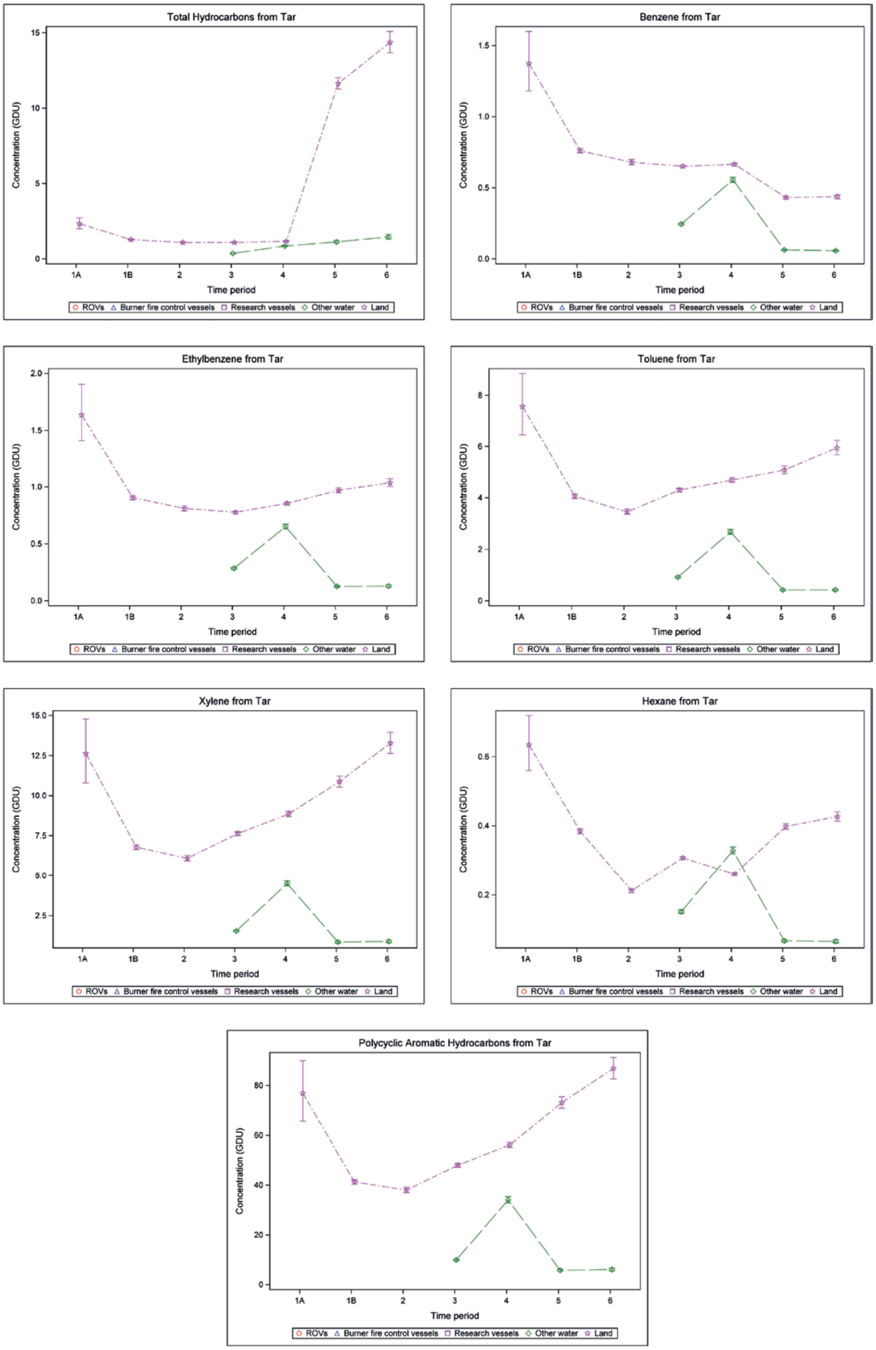


Figure 2. Modeled dermal exposure estimates of tar components (total hydrocarbons, benzene, ethylbenzene, toluene, xylene, *n*-hexane, and polycyclic aromatic hydrocarbons by broad activity groups by time period). Note that the scales (in GDUs) vary among the substances. ROVs = vessels piloting remotely operated vehicles. GDU = GuLF DREAM unit.

statistically higher exposures than the rig, ROV, burner fire control and RV workers, whether for oil or for tar (Fig. 1 and 2 and SM, Table S6). Although often statistically significant, the differences in the AMs for these workers were low and may not be meaningful.

For BTEX-H, the activities generally associated with the minimum values (<0.02 GDUs for all five chemicals) varied (SM, Table S6). Benzene AMs for oil and for tar across study participants reached 12.77 and 3.69 GDUs, respectively. AMs for ethylbenzene rose to 12.17 GDUs for oil and to 11.65 GDUs for tar. Similarly, the maximum of the toluene AMs from oil was 17.45 GDUs; for tar, the values rose to 42.37 GDUs. The maximums for xylene were 36.77 GDUs for oil and 88.18 GDUs for tar. The maximum AMs for *n*-hexane in oil were 2.22 GDUs and in tar, 5.56 GDUs. For PAHs, the maximum values were 219.31 GDUs for oil and 587.98 GDUs for tar.

Workers on the rig vessels, the ROV vessels and the burner fire control vessels were not considered exposed to dispersants (Table 2). The two components evaluated for dispersants were THC and xylene to represent petroleum distillates, hydrotreated light. The AM of the estimates for participants on land for THC were 25.41 and 18.88 GDUs and for xylene, 0.21 and 0.36 GDUs in TP1a and TP1b, respectively, the only time periods for which dispersants were used. For participants on the water the AMs were for THC, 80.37 and for xylene, 1.39 GDUs, respectively.

Correlations among the substances in oil across all time periods combined were generally high (Table 1). All substances had correlations of $\rho \geq 0.9$ with each other except *n*-hexane. The *n*-hexane correlation with both THC and toluene was $\rho = 0.87$, with ethylbenzene, rho = 0.84, xylene, $\rho = 0.83$ and with PAHs, $\rho = 0.82$. Correlations in tar remained high for ($\rho > 0.9$) for the relationships between toluene, xylene, *n*-hexane and PAHs. The relationships of THC and benzene, however, with each other and with the other substances of interest were lower ($\rho = 0.28$–0.86 for THC, $\rho = 0.70$–0.89 for benzene).

Discussion

We describe procedures for the development of dermal exposure estimates derived from an updated and enhanced version of DREAM (Gorman Ng *et al.*, 2021). We used participants' interview responses when available and JEM or industrial hygienist-entered values when not available. Even after deleting values <0.01 GDUs, minimum values were <0.05 GDUs for all substances in oil or tar. For THC, values ranged up to ~5 GDUs for oil and ~142 GDUs for tar. *n*-Hexane had GDU levels of the same order of magnitude as those of THC from oil, whereas the estimates for benzene, ethylbenzene, toluene, and xylene were substantially higher than THC for oil (up to 36.77 GDUs for xylene). For tar, exposures were lower for the BTEX-H chemicals. PAHs, in contrast, had the highest maximums of ~219 and ~588 GDUs for oil and tar, respectively. The broad ranges of estimates for the exposures suggest that there should be sufficient contrast to differentiate between low and high exposed individuals should these levels be associated with a health outcome.

There were statistically significant differences in the AM exposures to THC in oil over time for the other water workers after TP1b and for land workers after TP2. Similarly, for tar, other water workers' THC AMs rose after TP3, and land worker's THC AMs rose after TP4. We also evaluated THC exposure differences over broad groups of workers. Land workers had statistically higher THC AM exposures from oil than did all other workers in TP1a to TP4. A similar pattern was seen for tar: land workers had higher THC AMs from tar in TP3, TP4, TP5, and TP6 than other water workers. The fact that we found some differences suggests that we were able to, at least to some degree, increase the accuracy of our estimates by considering jobs/activities/tasks, location and time, thereby reducing misclassification error within the study participants.

It seems counterintuitive that exposures to oil and (for land workers) to tar should have risen over time. The model is complex, however, and has many variables that affect the estimates, making it difficult to identify the single reason for this finding. In the online SM, we explore the effect of concentration, vapor pressure and viscosity to learn how these varied by substance from 25% weathering (around TP2, generally reflecting lower exposures from oil) to 40% weathering (TP5 and 6, reflecting higher exposures to tar). Viscosity appeared to be the predominant variable for ethylbenzene, toluene, xylene and PAHs. Viscosity results in stickiness and the stickier the substance, the longer it will stay on the skin, thus increasing exposure. VP was the most important variable for THC and *n*-hexane. For both substances, the VP decreased substantially as the oil became more weathered, resulting in an increase in the VP score (the score being the inverse of VP) seen only with THC, benzene, and *n*-hexane, and therefore an increase in the exposure for THC and *n*-hexane. This decrease in VP and the increase in viscosity should have resulted in an increase in benzene exposures over time; however, concentration dominated the change seen in benzene by being the second largest decrease among the substances. See Stenzel, Arnold *et al.* (2021) for more details on weathering and its impact on the oil components. Less of a contrast between benzene, THC, and *n*-hexane may have been seen had the differences in the VP score categories more closely reflected the differences in the VP (i.e. for 0–40% weathering, THC VP changed from 3986

Table 1. Correlations (ρ) between total hydrocarbons, benzene, ethylbenzene, toluene, xylene, *n*-hexane and polycyclic aromatic hydrocarbons in oil and tar.

	Correlation (*ρ*) among oil-related components (*N* = 1643[a])						
	THC	**Benzene**	**Ethylbenzene**	**Toluene**	**Xylene**	**n-Hexane**	**PAHs**
THC	1.00	0.98	>0.99	>0.99	0.98	0.87	0.98
Benzene	0.98	1.00	0.97	0.97	0.95	0.93	0.95
Ethylbenzene			1.00	>0.99	0.98	0.84	0.98
Toluene				1.00	0.98	0.87	0.98
Xylene					1.00	0.83	>0.99
n-Hexane						1.00	0.82
PAHs							1.00

	Correlation (*ρ*) among tar-related components (*N* = 773)						
	THC	**Benzene**	**Ethylbenzene**	**Toluene**	**Xylene**	***n*-Hexane**	PAHs
THC	1.00	0.28	0.77	0.68	0.77	0.86	0.80
Benzene		1.00	0.83	0.89	0.82	0.70	0.79
Ethylbenzene			1.00	>0.99	>0.99	0.97	>0.99
Toluene				1.00	>0.99	0.94	0.98
Xylene					1.00	0.96	>0.99
n-Hexane						1.00	0.97
PAHs							1.00

THC: total hydrocarbons. PAHs: polycyclic aromatic hydrocarbons.
N = number of exposure groups.

Table 2. Estimates of dispersant exposures (in GDUs).[a]

Chemical	Broad group	TP1a	TP1b
		AM (LCL, UCL)	AM (LCL, UCL)
THC	All land	25.41 (22.32, 28.93)	18.88 (17.26, 20.64)
THC	All other water operations	NA[b]	80.37 (62.54, 103.30)
Xylene	All land	0.21 (0.19, 0.24)	0.36 (0.33, 0.38)
Xylene	All other water operations	NA	1.39 (1.08, 1.79)

GDUs: GuLF DREAM units (see text for definition.) TP1a = time period 1a (April 22–May 14, 2010). TP1b = time period 1b (May 15–July 15, 2010). AM (AMLC, AMUC) = arithmetic mean (AM lower confidence value, AM upper confidence value). THC = total hydrocarbons.

[a]Workers on rig vessels, vessels piloting remotely operated vehicles (ROVs), burner fire control vessels, and research vessels were not identified as having dispersant exposures.

[b]NA = not applicable. No water workers in TP1a reported having contact with dispersants.

to 821 Pa, respectively, whereas the scores changed from 0.01 to 0.1, respectively. For *n*-hexane, the respective values were 1025 to 36 Pa, with respective scores of 0.01 to 1.0. For benzene, in contrast, those same respective values were 115 to 29 Pa but scores of 0.1 to 1.)

A second reason for the rise in exposures is likely an artifact of the data, in that during the later time periods, higher exposed activities comprise a larger percentage of the activities being performed. We compared the TP2 median THC exposures of the activities performed in TP2 and in TP5 versus the median for activities only performed in TP2 and found a higher median THC exposure in the former (i.e. the first comparison of TP2 and TP5). In addition, the later activities were likely more associated with tar (with the higher viscosity) than with oil (with the lower viscosity).

The information used as inputs to the model was primarily from self-reports of the study participants for several reasons. We were able to observe only a limited number of operations, as almost all OSRC activities had ceased by the time the GuLF STUDY was initiated. In addition, there were thousands of workers across 4 states, and the oil was detected over more than 112 100 km^2 (69 656 mi^2) of water (Westerholm and Rauch, 2016). Moreover, although there was guidance on what protective equipment to wear for some of the activities, on-site industrial hygienists were responsible for modifying the recommended equipment given the specific work conditions encountered on any given day.

The estimates for the various oil-related substances generally were highly correlated ($\rho > 0.9$) for oil. In contrast, for tar the relationships between THC and benzene found lower correlations ($\rho = 0.3–0.9$), although the relationships among the other substances remained high. The high correlations in general are not surprising, however, because the variables associated with the reported job/activity/task were assigned the same weights regardless of the substance. The differences only came with substance-related variables, i.e. the concentration, vapor pressure score, and viscosity score. This finding may make it difficult to identify if a particular substance is uniquely associated with an adverse health effect.

It is not known how accurate the participants' reports were. First, we asked about skin/clothing contact with tar, defined in the interview as a "solid or gooey oily residue", and then with oil ("oil or oily water") to reduce the likelihood of participants confusing oil with tar. Some participants still, however, may have confused the two, particularly after the oil had undergone some weathering. This confusion primarily would have affected participants who worked on land, although participants performing water activities after TP3 (August 10–September 1, 2010) could also have been confused. Depending on the substance, this confusion could result in an over- or underestimation of exposures. Second, we asked about frequency of skin exposure in terms of the proportion of a day (none, less than half, about half, more than half, all of it), but we asked about hand exposure in terms of hours and minutes. In the home visit, we asked how often contact with "a chemical" occurred with each body part (<1 day per month, 1-4 days per month, 1–5 days per week or almost every day), and we assumed that the frequency on the body part would have been the same for all exposures of interest, which likely introduced some error in the exposure estimation. It is not clear what the magnitude would be and whether the error would result primarily in an over or underestimation of exposure. In contrast, reports of jobs/activities/tasks was likely to have been good. Teschke *et al.* (2002), found that reporting of job classifications generally had good agreement with records (Teschke *et al.*, 2002), and recollection was better for recent jobs than jobs further in the past. It is likely that reporting of activities in our study would be easier to recall than job classifications, particularly with the short time interval between the spill and the enrollment interview 1–3 years later and the unusual circumstances of the event.

We did not ask about changes over time. We know activities changed, as did the weathering of oil. While we made allowances for changes in weathering, we do not know if the specific tasks performed for a particular activity changed over time and had no way to adjust for such an occurrence if it had transpired. It is unlikely however, that the respondents would have been able to differentiate among exposure conditions in each of up to the seven time periods of the study. Furthermore, participants most often completed the enrollment questionnaire by telephone and the investigators were reluctant to add such detail to an already long interview. To reduce the potential misclassification associated with not obtaining participant information for the relevant time periods, we imputed missing data based on the responses from only those individuals who worked in each specific time period. This approach could have biased estimates in other time periods if participants tended to report the characteristics of the period with the greatest, or the least, skin burden.

Estimates for some of the variables were assigned by the study industrial hygienists. Of these, some were substance-specific data (concentration, vapor pressure score, viscosity score) that changed with the degree of weathering. Other variables were the percentages of the body parts exposed and the emission pathway, both of which we did not think that respondents would have been able to provide. We had no way to validate our assignments because the oil spill response and clean-up effort was largely completed by the time the exposure assessment started, and we were able to observe few workers during the performance of their jobs. We reviewed, however, the extensive amount of documentation, particularly on weathering (see the SM in Stenzel, Arnold *et al.*, 2021) and the substantial number of photographs taken of OSRC workers to obtain information on work activities.

We did not evaluate the day-to-day variability of exposures experienced by an individual due to the lack of day-to-day information. The model developed point estimates from over 50 exposure variables. Given the amount of uncertainty in the self-reports, this additional uncertainty was unlikely to have provided useful information.

In our evaluation of the GuLF DREAM model, we were unable to validate our actual estimates with measurements taken during the OSRC, and we were unable to find any useful dermal measurements taken on oil spill workers. The only oil spill dermal measurements we found were from a National Institute for Occupational Safety and Health (NIOSH) study of workers responding to the *Exxon Valdez* spill (Gorman *et al.*, 1991), which reported that dermal exposure levels were higher prior to the work shift than after the work shift. Thus, we used two published datasets: one, for oil, represented by a study measuring heavy fuel oil exposures in a variety of oil-using industries (Christopher *et al.* 2007, 2011) and the other, for tar, represented by a study of an asphalt paving operation (Cavallari *et al.*, 2012). Overall, the evaluation demonstrated a moderate correlation (ρ = 0.59) between GuLF DREAM exposure estimates and hand wash and wipe measurements. This correlation was somewhat lower than the corresponding correlation calculated for DREAM (ρ = 0.78) (van Wendel de Joode, 2005). This is not surprising because the DREAM validation by van Wendel de Joode (2005) involved exposure assessors who had often directly observed the workers during the exposure measurements. GuLF DREAM was evaluated with previously collected datasets of other investigators, so that study investigators did not have the opportunity to observe the measured workers directly. We were unable to evaluate the model for body parts other than hands due to insufficient measurements in the published studies. The original DREAM model found poorer correlation with other body parts.

Limitations of this work include dependence on the study participants' self-reports; our inability to observe the response and clean-up activities; lack of time-varying information on specific work; lack of information on variability; and lack of an evaluation on parts of the body other than hands. In addition, the reports of dispersants may have been overreported; positive responses suggest more participants had exposure than expected. Finally, we estimated exposure using a dimensionless unit and do not know how it relates to absorption.

This is, however, the first study to provide relative dermal exposure estimates from working on the response and clean-up of an oil spill under the many varied activities that took place during the OSRC. We used a deterministic model, an approach that has been used in other studies without measurement data (Vermeulen *et al.*, 2002). The evaluation, although limited, provides some information on both oil and asphalt (similar to tar) exposures. The dermal estimates show a different exposure pattern from inhalation, allowing investigators to investigate different routes of exposure and disease mechanisms. Finally, the estimates yielded large contrasts across the range of exposure levels for the substances of interest. This broad range, while likely to contain misclassification error, should allow investigators to examine risks associated with categories of exposure, such as low-, medium-, and high-exposed workers to distinguish between the most heavily and least exposed type of workers.

Conclusions

Assessment of dermal exposures is difficult, and in this study, there were no dermal measurements made during the response and clean-up operations. We used a deterministic model with participant-reported and industrial hygienist-supplied data. Dermal exposures to seven substances contained in oil and tar and to dispersants were estimated for each of the jobs/activities/tasks by location and time period reported by the nearly 24 000 workers in the study. The range of dermal estimates was substantial for many of the exposures of interest. This provides some confidence that there may be enough contrast to identify exposure-related differences in workers' health, should they exist; however, the high correlation among most of the exposures makes it unclear whether we will be able to identify a particular substance among the many assessed. Dermal exposure assessment is still in its infancy and the usefulness of deterministic models needs further evaluation.

Supplementary data

Supplementary data are available at *Annals of Work Exposures and Health* online.

Acknowledgments

We thank Wendy McDowell and Kaitlyn Rousch of McDowell Safety and Health Services, Inc. and Matthew Curry, Braxton Jackson, John McGrath and Kate Christenbury of Social & Scientific Systems, Inc. for the tremendous help they provided on this study. We also thank the workers for their participation in this study.

Funding

This study was funded by the NIH Common Fund and the Intramural Research Program of the National Institute of Health, National Institute of Environmental Health Sciences (ZO1 ES 102945).

Conflict of interest

Prof Cherrie is currently undertaking consulting work related to the *Deepwater Horizon* disaster. All of his involvement with this paper was prior to any potential conflict of interest arising.

Data availability

The data underlying this article will be shared on reasonable request, consistent with protections for the privacy of study participants and existing multi-party agreements. Requests should be made following instructions on the study website https://gulfstudy.nih.gov.

References

Arnold S, Stewart PA, Pratt GC *et al.* (2021) Estimation of aerosol concentrations of oil dispersants COREXIT™ EC9527a and EC9500A during the *Deepwater Horizon* oil spill response and clean-up operations. *Ann Work Expo Health*; **66**: i188–i201.

Cavallari JM, Osborn LV, Snawder JE *et al.* (2012) Predictors of dermal exposures to polycyclic aromatic compounds among hot-mix asphalt paving workers. *Ann Occup Hyg*; **56**: 125–37.

Christopher Y, van Tongeren M, Cowie H *et al.* (2007) Occupational dermal exposure to heavy fuel oils. Research Report TM/07/05. Edinburgh, UK: IOM.

Christopher Y, Van Tongeren M, Urbanus J *et al.* (2011) An assessment of dermal exposure to heavy fuel oil (HFO) in occupational settings. *Ann Occup Hyg*; **55**: 319–28.

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

Gorman Ng M, Cherrie JW, Sleeuwenhoek A *et al.* (2021). GuLF DREAM: a model to estimate dermal exposure among oil spill response and clean-up workers. *Ann Work Expo Health*; **66**: i218–i233.

Groth CP, Banerjee S, Ramachandran G *et al.* (2021) Methods for the analysis of 26 million VOC area measurements during the *Deepwater Horizon* oil spill clean-up. *Ann Work Expo Health*; **66**: i140–i155.

Groth CP, Huynh TB, Banerjee S *et al.* (2021) Linear relationships between total hydrocarbons and benzene, toluene, ethylbenzene, xylene, and *n*-hexane during the *Deepwater Horizon* response and clean-up. *Ann Work Expo Health*; **66**: i71–i88.

Huynh TB, Groth CP, Ramachandran G *et al.* (2021a) Estimates of occupational inhalation exposures to six oil-related compounds on the four rig vessels responding to the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; **66**: i89–i110.

Huynh TB, Groth CP, Ramachandran G *et al.* (2021b) Estimates of inhalation exposures to oil-related components on the supporting vessels during the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; **66**: i111–i123.

Huynh TB, Groth CP, Ramachandran G *et al.* (2021c) Estimates of inhalation exposures among land workers during the *Deepwater Horizon* oil spill clean-up operations. *Ann Work Expo Health*; **66**: i124–i139.

Kwok RK, Engel LS, Miller AK *et al.*; GuLF STUDY Research Team. (2017) The GuLF STUDY: a prospective study of persons involved in the *deepwater horizon* oil spill response and clean-up. *Environ Health Perspect*; **125**: 570–8.

NIOSH. (2011) NIOSH *Deepwater Horizon* Roster Summary Report. Available at https://www.cdc.gov/niosh/updates/upd-12-19-11.html. (accessed 16 June 2018).

NIOSH. (1994) Pocket Guide to Chemical Hazards. DHHS (NIOSH) Publication No. 94-116.

Pratt GC, Stenzel MR, Kwok RK *et al.* (2021) Modeled air pollution from *in situ* burning and flaring of oil and gas released following the *Deepwater Horizon* disaster. *Ann Work Expo Health*; **66**: i172–i187.

Ramachandran G, Groth CP, Huynh TB *et al.* (2021) Using real-time area VOC measurements to estimate total hydrocarbons exposures to workers involved in the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; **66**: i156–i171.

Stenzel MR, Arnold SF, Ramachandran G *et al.* (2021) Estimation of airborne concentrations of oil dispersants COREXIT™ EC9527A and EC9500A, volatile components associated with the *Deepwater Horizon* oil spill response and clean-up operations. *Ann Work Expo Health*; **66**: i202–i217.

Stenzel MR, Groth CP, Huynh TB *et al.* (2021) Exposure group development in support of the NIEHS GuLF STUDY. *Ann Work Expo Health*; **66**: i23–i55.

Stewart P, Groth C, Huynh TB *et al.* (2021) Assessing exposures from the *Deepwater Horizon* oil spill response and clean-up. *Ann Work Expo Health*; **66**; i3–i22.

Teschke K, Olshan AF, Daniels JL *et al.* (2002) Occupational exposure assessment in case–control studies: opportunities for improvement. *Occup Environ Med*; **59**: 575–594.

Van Wendel de Joode B, Brouwer DH, Vermeulen R *et al.* (2003) DREAM: a method for semi-quantitative dermal exposure assessment. *Ann Occup Hyg*; **47**: 71–87.

Van Wendel de Joode B, Vermeulen R, van Hemmen JJ *et al.* (2005) Accuracy of a semiquantitative method for Dermal Exposure Assessment (DREAM). *Occup Environ Med*; **62**: 623–32.

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

Westerholm DA, Rauch SD III. (2016) *Deepwater Horizon* oil spill: final programmatic damage assessment and restoration plan and final programmatic environmental impact statement. https://repository.library.noaa.gov/view/noaa/18084 (accessed 6 June 2020)