Annals of Work Exposures and Health, 2022, Vol. 66, No. S1, i156–i171
https://doi.org/10.1093/annweh/wxab066
Advance Access publication 13 September 2021
Original Article

Original Article

Using Real-Time Area VOC Measurements to Estimate Total Hydrocarbons Exposures to Workers Involved in the *Deepwater Horizon* Oil Spill

Gurumurthy Ramachandran[1,*], Caroline P. Groth[2], Tran B. Huynh[3], Sudipto Banerjee[4], Patricia A. Stewart[5], Lawrence S. Engel[6,7], Richard K. Kwok[6,8], Dale P. Sandler[6] and Mark Stenzel[9]

[1]Department of Environmental Health and Engineering, Bloomberg School of Public Health, Johns Hopkins University, 615 N Wolfe Street, Baltimore, MD 21205, USA; [2]Department of Epidemiology and Biostatistics, WVU School of Public Health, West Virginia University, One Medical Center Drive, Morgantown, WV 26506, USA; [3]Department of Environmental and Occupational Health, Dornsife School of Public Health, Drexel University, 3215 Market Street, Philadelphia, PA 19104, USA; [4]Department of Biostatistics, UCLA Fielding School of Public Health, University of California—Los Angeles, 650 Charles E. Young Drive, Los Angeles, CA 90095, USA; [5]Stewart Exposure Assessments, LLC, 6045 N 27th Street, Arlington, VA 22207, USA; [6]Epidemiology Branch, National Institute of Environmental Health Sciences, 111 T.W. Alexander Drive—MD A3-05, Research Triangle Park, NC 27709, USA; [7]Department of Epidemiology, University of North Carolina at Chapel Hill, 35 Dauer Drive, Chapel Hill, NC 27599, USA; [8]Office of the Director, National Institute of Environmental Health Sciences, 9000 Rockville Pike, Bethesda, MD 20892, USA; [9]Exposure Assessment Applications, LLC, 6045 N 27th Street, Arlington, VA 22207, USA

*Author to whom correspondence should be addressed. Tel: +1-612-619-6702; e-mail: gramach5@jhu.edu

Submitted 17 October 2020; revised 30 July 2021; editorial decision 31 July 2021; revised version accepted 2 August 2021.

Abstract

Even though the *Deepwater Horizon* oil spill response and clean-up (OSRC) had one of the largest exposure monitoring efforts of any oil spill, a number of exposure groups did not have sufficient personal data available or there were gaps in days measured to adequately characterize exposures for the GuLF STUDY, an epidemiologic study investigating the health of the OSRC workers. Area measurements were available from real-time air monitoring instruments and used to supplement the personal exposure measurements.

Objectives: The objective was to present a method that used real-time volatile organic compounds (VOCs) area measurements transformed to daily total hydrocarbons (THC) time-weighted averages (TWAs) to supplement THC personal full-shift measurements collected using passive charcoal badges. A second objective was to develop exposure statistics using these data for workers on vessels piloting remotely operated vehicle (ROV) vessels and other marine vessels (MVs) not at the job title level, but at the vessel level.

Methods: From hourly vessel averages derived from ~26 million real-time VOC measurements, we estimated full-shift VOC TWAs. Then, we determined the relationship between these TWAs and corresponding full-shift THC personal measurements taken on the same vessel-day. We used this relationship to convert the full-shift VOC measurements to full-shift 'THC' TWA estimates when no personal THC measurements existed on a vessel-day. We then calculated arithmetic means (AMs) and other statistics of THC exposures for each vessel.

Results: The VOC-derived estimates substantially supplemented the THC personal measurements, with the number of vessel-days for which we have exposure estimates increasing by ~60%. The estimates of the AMs are some of the highest observed in the GuLF STUDY. As expected, the AMs decreased over time, consistent with our findings on other vessels.

Conclusions: Despite the inherent limitations of using real-time area measurements, we were able to develop additional daily observations of personal THC exposures for workers on the ROV vessels and other MVs over time. The estimates likely resulted in more representative estimates of the AMs in the GuLF STUDY. The method used here can be applied in other occupational settings and industries for personal exposure estimation where large amounts of area measurements and more limited numbers of personal measurements are available.

Keywords: *Deepwater Horizon*; linear regression; real-time area measurements; total hydrocarbons; volatile organic compounds

Introduction

Approximately 55 000 workers were rostered as having possible involvement in the oil spill response and clean-up (OSRC) activities that occurred following the 20 April 2010 *Deepwater Horizon* (*DWH*) oil drilling rig explosion in the Gulf of Mexico (NIOSH, 2011). The explosion resulted in a release of crude oil, and many volatile chemicals were released into the air from this oil as it rose to the water surface over the next few months. Workers were exposed to these chemicals at levels that varied by exposure determinants including their vessel, job or activity, location (area of the Gulf or state), exposure duration, and dates worked relative to important study events. The National Institute for Environmental Health Science's Long-term Follow-up Study (GuLF STUDY) has enrolled 32 608 study participants comprising 24 937 workers and 7671 individuals who completed training but were not hired. The primary goals of the study are (i) to develop airborne exposure estimates particularly to total hydrocarbons (THC, measured as total petroleum hydrocarbons) and benzene, toluene, ethylbenzene, xylene, and *n*-hexane (BTEX-H) for all study participants (Stewart, Groth *et al.*, 2021); (ii) investigate if there is an association between the exposures and detrimental health outcomes (Kwok *et al.*, 2017). To carry out the first goal, two large databases were used: one of full-shift personal air measurements of THC and the BTEX-H chemicals and a second of real-time area measurements for volatile organic compounds (VOCs).

The first objective of this paper is to, for calculation of exposure statistics, develop a set of VOC estimates comparable to full-shift THC personal measurements to provide observations on vessel-days when THC measurements were unavailable. We evaluated the relationship between full-shift VOC time-weighted averages (TWAs) obtained from ~26 million area direct-reading VOC measurements (Groth, Banerjee *et al.*, 2021) and THC personal full-shift measurements taken on the same vessel-day. We then used this relationship to convert the VOC TWAs to 'THC' TWA estimates for use when no personal THC measurements existed on a given vessel-day. The second objective was to calculate arithmetic means (AMs) and other statistics to describe exposures of the GuLF STUDY workers on the two types of vessels on which the VOC measurements were collected, i.e. vessels piloting remotely operated vehicle (ROV) vessels and other marine vessels (MVs). Unlike for the workers on the rig ships (Huynh *et al.*, 2021b), we do not have information on participants' jobs on the ROV vessels and the other MVs; therefore, exposures were developed at the vessel level.

Stewart, Groth *et al.* (2021) describe the overall exposure assessment for the study, and Stenzel, Groth, Banerjee *et al.* (2021) describe the recalculation of the exposure measurements to reduce THC censoring from 83 to 11%. Inhalation exposures to THC and BTEX-H of other workers performing OSRC-related activities (Huynh *et al.*, 2021a, b, c) and to other substances [i.e. aerosols and vapors from dispersants (Arnold *et al.*, 2021; Stenzel, Arnold, *et al.*, 2021, respectively)], $PM_{2.5}$ (Pratt *et al.*, 2021), and oil mist (Stewart, Groth, *et al.*, 2021), as well as dermal exposures (Gorman Ng *et al.*, 2021; Stewart, Gorman Ng, *et al.*, 2021) are presented elsewhere.

Background

Vessels piloting remotely operated vehicles

The ROV vessels were MVs that used tethered ROVs to inspect damage of underwater equipment; perform repair work; assist in the installation and removal of equipment; provide video to assist in positioning equipment for capturing oil or mechanically capping the well by the replacement of the damaged blowout preventer (BOP, used to prevent an uncontrolled oil release) and well casing; remotely activate the valve to block the BOP; photograph oil leaks and underwater wildlife; observe the quantity of oil on the Gulf floor; install acoustic transponders; and hold the subsea dispersant injection spray boom into the underwater oil plume. At a depth of 1524 m (5000 ft) below the water surface, the ROVs had to overcoming harsh pressure and temperature conditions and low visibility. At peak, 27 ROVs were deployed from the 14 ROV vessels. Work was transferred among ROV vessels as needed to maintain continuity and advance time-critical operations. In addition to the workers who piloted the ROVs, there were workers who performed maintenance and cleaned the ROVs of oil once brought back up on deck, before the ROVs were redeployed. These vessels generally spent most of their time near the wellhead where THC measurements were among the highest of the OSRC operations.

Other marine vessels

Our category of other MVs contained various types of vessels, including specialized 'stimulation ships' that were specifically adapted to pumping fluids into the well under high pressure for the well closure attempts; vessels working in support of the drill rig ships; semisubmersible deep-water drilling platform ships; deep-water supply vessels; vessels spraying dispersant onto the water's surface; fire control vessels spraying water to cool vessel surfaces; and service vessels fitted for towing and anchor handling. Real-time VOC measurements were available on 21 MVs. Most of these MVs spent substantial portions of their time at the wellhead.

Time periods of interest

The *DWH* exploded on 20 April 2010 and sank on 22 April 2010, resulting in damage to the underwater equipment, including severing of the riser pipe that connected the *DWH* and the wellhead, resulting in the release of the crude oil. Various MVs arrived at the scene rapidly and started search and rescue of workers, firefighting operations, and attempts to close the BOP over the well. The BOP, however, was damaged and could not be closed. The four rig vessels (described in Huynh *et al.*, 2021b) arrived within a couple of weeks of the explosion to stop the oil release and to drill two new wells and were supported in their operations by the ROV vessels and MVs that are the focus of this paper. For the GuLF STUDY, we divided the response operations into time periods that were expected to result in changes in exposures, and we assigned exposure estimates to all OSRC study participants on these vessels based on the average of the measurements within each of these time periods. Only events affecting the ROV vessels and MVs are described below.

Time period 1a (TP1a) (22 April to 14 May 2010). Oil was flowing from the well into the Gulf of Mexico. Attempts to stop the flow by activating the BOP (supported by the ROV vessels and a small number of MVs) were unsuccessful. Limited testing of subsea injection (1524 m below the water surface) of dispersants was done by one ROV vessel.

Time period 1b (TP1b) (15 May to 15 July 2010). Oil continued to flow from the well. All 14 ROV vessels and most of the 21 MVs in this report supported the four rigs that were either trying to stop the oil release or drilling either of two relief wells in the wellhead area. One ROV vessel routinely injected dispersant 1524 m below the water surface starting 15 May, and up to five MVs sprayed dispersant on the water surface. Three fire control MVs sprayed cooling water onto two of the rig vessels and one MV that was flaring oil. The well was successfully mechanically capped on 15 July by replacing the damaged BOP. From this point on, little new oil was released from the BOP.

Time period 2 (TP2) (16 July to 10 August 2010). On 4 August, the 'static kill' of the well was accomplished (i.e. 'heavy mud' was pumped into the top ~1524 m of the damaged well casing followed, a day later, by pumping cement, which relieved pressure on the well). Most of the ROV vessels and MVs were still supporting the operations. By 10 August, after several days of testing, the effort was deemed a success.

Time period 3 (TP3) (11 August to 30 September 2010). On 16 September, the first relief well intersected the original well ~18 000 ft below the Gulf floor. Pumping of cement into the base of the new well began, and the well was declared successfully sealed on 19 September. Most of the ROV vessels continued operations to remove or reposition the underwater equipment, while the MVs started leaving the area.

Time period 4 (TP4) (1 October 2010 to 31 December 2010). Decontamination and refurbishing of the vessels may have occurred on any remaining vessels but no information or measurements are available. All remaining ROV vessels and MVs left the area.

Personal measurement database

The responsible party (RP) of the spill (as designated by the US government) used organic vapor badge passive dosimeters (3M 3500 or 3520, or Assay Technology 521) to collect ~28 000 personal air samples analyzed for multiple analytes, including THC and BTEX-H, resulting in over 160 000 individual measurements (Stenzel, Groth, Banerjee *et al.*, 2021). A database with these measurements also contained information that was used to develop and characterize unique job or work groups [vessel, job title (where applicable) or activity; area of the Gulf/US state; and time period], which allowed us to pair measurements to develop exposure estimates for THC and BTEX-H to these unique job/work groups (Stenzel, Groth, Huynh *et al.*, 2021) to link to study participants.

The original measurements received by the RP had an unexpectantly high percentage of censored (below the limit of detection, LOD) data. Typically to determine the true LOD of an analytic method, labs prepare standards at various concentrations of an analyte, and a calibration curve is developed between the concentrations of the standards and the area under the curve of the gas chromatograph readout. The true LOD of the analytical procedure, in micrograms (µg), corresponds to an area under the curve that is three times the average area of the blank samples. This LOD is based on the capability of the method, i.e. the lowest concentration in the sample that can be quantified with the desired precision and accuracy and can be distinguished from the background concentrations observed in blank samples. However, the labs used by the RP to analyze the samples were focused on assessing compliance. The range of concentrations expected when compliance is of interest is generally 0.10–2 times the target concentration, such as the occupational exposure limit (OEL). The labs reported the concentration that corresponded to the lowest calibration standard as the LOD. Thus, although the labs referred to their censored data as being at or below the LOD, in actual practice they reported a measurement as censored if the result was below their calibration range. This reported LOD (RLOD) may be distinct from the analytic method's LOD. Our original review of the measurements found that the RLODs were much higher than that would have been expected based on the published expected performance of the analytical methods. The RP requested that both labs recalculate the measurement data using the analytic method's LOD to reflect the analytical capability of the methods. The recalculation of the data did not involve re-analyzing the samples, but rather simply recalculating the measurement results using the labs' same calibration curves as they had used earlier.

Measurements were available on many of the over 9000 vessels deployed in some capacity: four rig ships (Huynh *et al.*, 2021b), ROV vessels, other MVs, and a large number of smaller vessels (Huynh *et al.*, 2021a). In addition, measurements covered operations on land (Huynh *et al.*, 2021c). For some of these vessels (e.g. the rig ships), the measurements were sufficient for estimating personal exposures for most job groups (Huynh *et al.*, 2021b). However, the coverage of days when personal measurements were collected was more limited on other vessels, particularly on the ROV vessels and MVs that worked near the wellhead.

Real-time VOC measurements database

The second database contained results of >26 million VOC direct-reading area measurements of ~1-min duration collected at various stationary locations on 38 vessels from 30 April 2010 through 29 August 2010. The RP collected these measurements to monitor (i) the air around the operation activities to protect potential downwind receptors; (ii) air in the vicinity of operations activities to protect worker health; and (iii) specific activities to support safe operations (BP Gulf Science Data, 2016). These measurements were used by the RP to support decisions that had to be made on short-term bases, such as providing respiratory protection or initiating efforts to suppress the oil vapors emanating from the leaking oil rising to the water surface.

Of the 14 ROV vessels and an unknown number of MVs involved in the OSRC efforts, acceptable VOC measurements were taken on 13 ROV vessels and 21 MVs. The direct-reading instruments were AreaRae and MultiRae (RAE Systems Instruments) portable multi-gas monitors that had a photoionization detector equipped with a 10.6-eV UV lamp to detect VOC concentrations. The manufacturer reported that the instruments had a 10-s response time and could be set to collect measurements from seconds to minutes. The instruments stored measurements to one decimal place, at a frequency of 1 min. Some of these measurements were taken on the same vessel-day as the THC personal measurements; others were not. More than one instrument may have been sampling at any one time on any vessel (median number of instruments per vessel = 7); however, the locations of the instruments on each vessel were not recorded. The specific guidance the RP provided to the Air Monitoring Technicians regarding site monitoring locations was 'Vessel operators will work with the Air Monitoring Technicians to select real-time monitoring locations in common work areas and inside crew quarters. Additional monitors may be placed near the edge of the vessel or in other areas of interest, such as moon pools, to gain early indications of rising LEL levels' (lower explosive limit (LEL) (BP Gulf Science Data, 2016). See Groth, Banerjee *et al.* (2021) for more information on these measurements.

Methods

Fig. 1 presents an overview of our methods. Steps 1 and 2 of the figure describe how we developed from the ~26 000 million ~1-min duration VOC measurements VOC instrument-hour averages (instrument-hour TWAs) and then hourly averages by vessel, which is described in more detail in Groth, Banerjee *et al.* (2021). The results of that study were used in the work described in this paper (Fig. 1, steps 3–7).

Steps 1 and 2: development of VOC instrument-hour averages and hourly averages by vessel

A Bayesian analysis of variance (ANOVA; intercept only regression) that accounted for censoring was used to estimate 10 000 posterior samples of the geometric mean (GM) for each instrument-hour and the hourly geometric standard deviation (GSD) across all instruments in each hour, which allowed an estimate of uncertainty. From the GM and GSD posterior samples, 10 000 AMs were calculated using

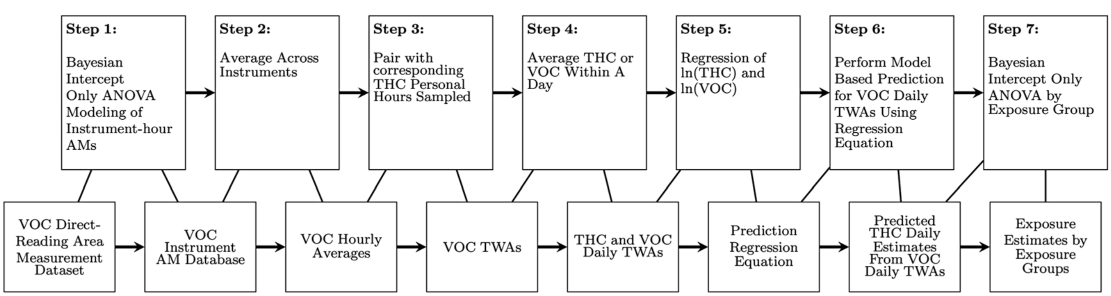
$$\mathrm{AM} = \mathrm{GM} \times e^{(1/2)(\log(\mathrm{GSD}))^2} \tag{1}$$

where e is the natural logarithm, to reflect the mean air concentration in each measured hour of an instrument's day on each vessel. The results of this analysis (the 2.5th, 97.5th, and median posterior AM estimates of each instrument-hour TWA) were compiled into an instrument-hour AM database for analysis. We chose to use the median estimate of the instrument-hour AMs in further calculations to reflect the true center of the distribution.

Finally, we further averaged across all instruments in an hour on a vessel-day to form 21 900 hourly averages.

Step 3: linking of VOC and THC measurements

We compiled a database that included the 21 900 VOC hourly averages (Groth, Banerjee *et al.*, 2021) and the non-censored (i.e. >LOD) THC personal samples of 4–18 h in duration taken on the same vessel-day (Fig. 1, step 3). Not all THC personal measurements were linked to a set of VOC hourly averages and conversely, VOC hourly averages were sometimes available for days when a THC measurement was not available. In both situations, those data were not included in the linking database, although VOC values not linked to THC measurements were used later (see Step 6: prediction of THC exposures from VOC daily averages on additional days). Data from two of the MVs were removed from this database because all the VOC measurements collected on those vessels represented instrument testing (calibration or bump testing using a gas standard of known concentration) or the instrument indicated (via a warning message) that it was not operating within the instrumentation performance criteria specifications.



VOC TWA: time weighted average over hours of THC personal sample; Daily TWA: time weighted average over a day; VOC: volatile organic compounds; THC: total hydrocarbons

Figure 1. Method for estimating arithmetic means of THC using THC measurements and VOC-derived "THC" TWA estimates. Groth, Banerjee *et al.* (2021) describe steps 1 and 2. This paper describes steps 3–7.

Development of VOC TWAs (step 3 continued)

On each day when both VOC hourly averages and THC full-shift measurements were available, we compared the hours over which each THC measurement was collected to the hours the VOC hourly averages covered. To be considered a pair for any given vessel-day, at least four VOC hourly averages had to overlap at least 4 h of the THC sampling period on the same vessel-day [4 h was the minimum sampling duration used to estimate exposures from THC measurements in other analyses (Huynh *et al*., 2021a,b,c)]. We averaged all VOC hourly averages that corresponded to the hours sampled by their paired personal THC measurement and referred to that average as the VOC 'TWA' in Fig. 1, step 3. As multiple THC personal measurements were taken on different people in the same day, VOC hourly averages were typically linked to more than one THC personal measurement. For example, if on a specific vessel one personal THC measurement was collected from 8:00 to 20:00 and a second personal THC measurement was collected from 9:00 to 15:00 on the same day, we calculated a VOC TWA corresponding to all available measured hours (but at least four) of the 12 h of the first THC measurement and a second VOC TWA corresponding to all available hours (but at least four) of the measured 6 h of the second THC measurement. In this case, the same (at least four) hourly VOC averages would have been used for the 9:00 to 15:00 period, but the 8:00 to 9:00 of the first THC sample and the 15:00 to 20:00 h of the second THC sample may also have had VOC hourly averages not paired to the other THC sample.

A series of simulation studies were performed to better understand the impact of including only non-censored THC measurements in the regression. We performed a series of Bayesian regressions on the VOC 'THC' TWAs (Step 4: calculation of THC and VOC daily averages) paired with corresponding THC personal samples including and not including the censored THC measurements. We accounted for censored measurements using the methods in Groth et al. (2017, 2018). Little difference in the regression coefficients was observed between these regressions, and therefore, we chose to use only non-censored THC observations (and their corresponding VOC TWAs) as this approach was simpler. Further details can be found in Supplementary Material (SM), document 1 (available at *Annals of Work Exposures and Health* online).

By averaging the hourly averages within a day, we eliminated information about hour-to-hour variability. In order to better assess the variability within a day, we performed an ANOVA assessing between and within vessel-day variances of the VOC data for the ROV vessels and MVs separately.

Step 4: calculation of THC and VOC daily averages

For each vessel, we calculated the average of all (non-censored) THC measurements that had a start time on each day of interest as well as the average of all corresponding VOC TWAs to obtain matched THC and VOC daily averages (Fig. 1, step 4). To ensure robustness of the average, we required at least three THC non-censored measurements on a given day and at least 3 days of non-censored THC measurements on a vessel within a time period for comparison. These criteria for the minimum number of measurements per day and per time period for a vessel were identified to increase the likelihood that the measurements were representative of the air concentrations experienced by various workers located throughout the vessel over the day and over the time period, respectively. Analyses evaluating the impact of fewer sets of VOC:THC matches required for the daily average are provided in SM, document 1 (available at *Annals of Work Exposures and Health* online).

Step 5: derivation of the VOC:THC relationship

Using the paired daily averages of non-censored THC and of VOC TWAs, we regressed the natural logarithms of the THC and VOC averages (in ppb) against each other. The regression used for prediction of observation i can be described as

$$\ln(\mathrm{THC}_i) = \beta_0 + \beta_1 \times \ln(\mathrm{VOC}_i) + \varepsilon_i \qquad (2)$$

where β_0 is the intercept, β_1 is the slope, and ε_i is the error, which is independently and identically distributed $N(0, \sigma^2)$. We used a Bayesian model with weakly informative priors on the regression coefficients and variances (i.e. wide normal distributions centered at 0 for the regression coefficients and inverse gamma priors on the variance components). Weakly informative priors provide limited information to allow the data to drive the inference.

To determine the best relationship for predicting THC exposure from VOC exposure, we considered different relationships, i.e. specific vessels (where possible), vessel type, and time period, for both the ROV vessels and the MVs. After evaluation of the regressions, we chose the relationship of all combined ROV vessels in time periods 1a–1b as the prediction equation for both types of vessels for all time periods. Further details of this decision are provided in SM, document 1 (available at *Annals of Work Exposures and Health* online).

Step 6: prediction of THC exposures from VOC daily averages on additional days

We then identified vessel-days when VOC measurements were available, but a THC measurement was not. One ROV vessel was eliminated at this step because it had no VOC daily averages without corresponding THC matches. For each of these vessel-days, we averaged the (at least 4) VOC hourly averages on the vessel for that day to represent a full-shift TWA. Then, using a simple linear Bayesian prediction and equation (2) (with the natural log of the VOC average on the ppb scale to allow all results to be positive on the log scale), we predicted a full-shift 'THC' TWA estimate. We ran the model for 25 000 Markov Chain Monte Carlo iterations, generating 25 000 posterior estimates of the GMs and GSDs for each predicted daily 'THC' TWA estimate (see Groth, Huynh, *et al*., 2021, for additional details of this method), from which we calculated 25 000 posterior samples of each AM using equation (1). While other statistics and quantiles were developed, we used the median posterior AM estimate as the final prediction value for each observation.

Step 7: developing exposure estimates for ROV vessels and marine vessels

To generate the final exposure estimates for the work groups of interest, we applied a Bayesian intercept only ANOVA model accounting for censoring using the original THC full-shift measurements (including the censored ones) and the 'THC' TWA estimates combined. This model, as described in Huynh *et al.* (2021b), models the mean of ln(THC) using an intercept term. We placed uniform priors on the ln(GM) and the ln(GSD) based on previous GuLF STUDY information [ln(GM) between ln(0.025) and ln(50); ln(GSD) between ln(1.01) and ln(12)]. From each model, we obtained posterior estimates of the AM (derived from the GM and GSD), GM, GSD, and 95th percentiles. Further details of this modeling strategy are provided in the supplementary materials of Stewart, Groth *et al.* (2021). All analyses for all steps above were performed in rjags (Plummer, 2003, 2016) and R (R Core Team, 2017).

Results

Results of two additional MVs (making the number of MVs = 17) are omitted here because they did not have corresponding THC measurements to compare too.

The posterior estimates of β_0, β_1, the correlation (ρ), and the R^2 for equation (2) for TP1a–1b are provided in Table 1. The 95% credible intervals (Bayesian uncertainty intervals) around the slope estimates were tight

Table 1. Parameters for the ROV TP1a–1b regression. This relationship was used to predict VOC-derived 'THC' TWA estimates in all time periods on both ROV vessels and MVs. The median and 95% credible intervals (CI) are reported for the intercept, slope, and correlation. The median posterior R^2 estimate is also reported.

Parameter	Median	95% CI
Intercept (β_0)	3.17	(2.59, 3.76)
Slope (β_1)	0.63	(0.54, 0.71)
Correlation, ρ	0.78	(0.71, 0.84)
R^2	0.61	

(Table 1), suggesting relatively little variability around the slope. The VOC:THC relationship had a ρ = 0.78 (R^2 = 0.61) (Supplementary Fig. S1, document 1, available at *Annals of Work Exposures and Health* online). Further details of how we chose this relationship for the MVs and other time periods are in Supplementary Material, document 1 (available at *Annals of Work Exposures and Health* online).

Fig. 2 shows a summary of the days for which each type of information (daily averages derived only from THC measurements and daily VOC-derived 'THC' TWA estimates) was available for each ROV vessel and MV. The figure also shows the days for which THC measurements overlapped with the VOC measurements on the ROV vessels that were used to derive the prediction equation. The improvement in coverage varied by vessel, ranging from 0 to 3800% (Table 2). For example, the *Adriatic* had only 1 vessel-day coverage from THC measurements; the VOC data increased the overall coverage on that vessel to 38-day observations, 19 of which were in TP1a–1b, the period in which we saw the highest exposure levels. Other vessels such as the *Normand Commander* and the *Seacor Venture* had modest improvements in coverage (13 and 8.6%, respectively) as a result of the VOC predictions. Thus, the number of vessel-days for which we have daily exposure estimates (TP1a–1b and TP2–3) increased for all ROV vessels and MVs from 894 THC vessel-days to 1389 vessel-days (the latter comprising both the THC measurements and the VOC 'THC' TWA predictions), a 56.6% increase. The number increased from 551 to 918 in TP1a–1b.

Supplementary Fig. S2, document 1 (available at *Annals of Work Exposures and Health* online), shows an example of one vessel, the *Ocean Intervention III*, where the use of the VOC measurements improved the coverage of the days for which 'THC' observations were available. The top and bottom graphs show the days for which VOC and THC daily averages were available, respectively. The bottom graph also shows the days for which we predicted additional 'THC' TWA estimates from VOC data. There are a substantial number of days when both VOC and THC information was available, and the VOC and THC levels are correlated (Pearson correlation, $\rho = 0.82$).

We calculated for each set of measurements (THC or VOC 'THC' TWAs) the median and 95% credible intervals (uncertainty intervals in the Bayesian context) of all measurements for the AMs for all ROV vessels and all MVs by time period (Fig. 3). To judge significance in the Bayesian context (here referred to as a notable or credible difference), we observed if the 95% credible intervals for independent groups overlapped (non-overlapping intervals suggest statistical significance). Only all non-censored VOC-predicted daily 'THC' TWA estimates were considered in this analysis. The first bar in the figure represents an estimate only from the THC measurements. The second bar represents the 'THC' TWA estimates predicted from the VOC measurements. The third bar represents the combined THC and the 'THC' TWA estimates (because the third bar incorporates both measurements, this bar cannot be statistically compared to the first two). The 'THC' TWA estimates represent a different data source than the THC personal measurements; therefore, they are not expected to, and should not necessarily, generate identical estimates of exposure. This graph shows that the 'THC' TWA data were notably different than the THC measurements for the ROV vessels in TP1a–1b and in TP2–3, as demonstrated by the non-overlapping credible intervals. This was not the case for MVs, where there was no credible difference between the 'THC' TWAs and THC measurements either in TP1a–1b or in TP2–3. For the ROV vessels, the figure shows that there was a credible difference between the AMs of TP1a–1b and TP2–3; for the MVs, there was a credible difference in the AMs of TP1a–1b and TP2–3. Considering the combined THC and the 'THC' TWA estimates, there is a credible difference between the AMs of the ROV vessels and the MVs for each time period.

Table 3 displays the AM, GM, and GSD median posterior estimates generated from the combined THC personal samples and the VOC-derived 'THC' TWA estimates by vessel for TP1a–1b and TP2–3. The overall AM exposures substantially declined on the ROV vessels from 5.35 ppm in TP1a–1b to 0.54 ppm in TP2–3 and on the MVs from 2.25 to 1.03 ppm over the same time periods. Further statistics are provided in SM, document

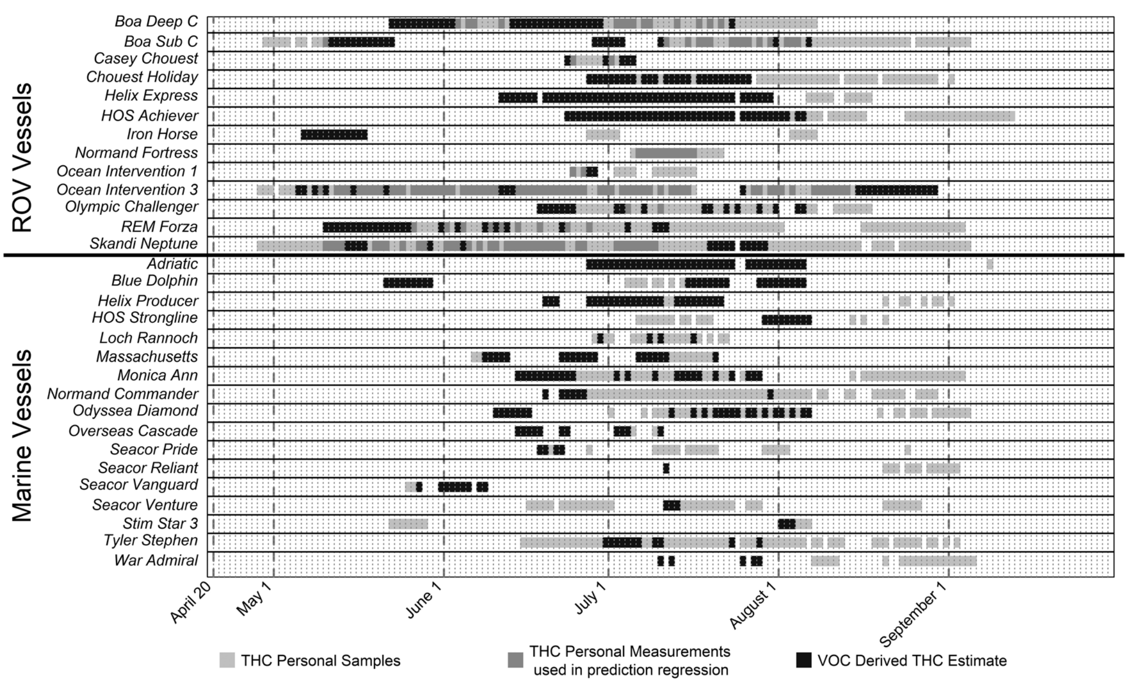


Figure 2. Days for which THC and VOC measurements were available for 13 ROV vessels and 17 MVs. The figure shows days for which only THC personal samples were available (the light squares), days when both VOC and THC measurements were available and used in the regression analysis (darker squares), and days when only VOC measurements were available for the VOC-derived "THC" TWA estimates (darkest squares).

Table 2. The number of days for which only THC personal samples were available; both VOs and THC measurements were available for development of the prediction equation; VOC-derived 'THC' TWA estimates were developed; and either a THC personal sample or a VOC-derived 'THC' TWA estimate was available. The last column identifies the percentage increase in the number of available measurements for estimating AMs for each ROV vessel and MV.

Type of vessel	Vessel name	N days with THC personal samples	*N* days used for regression of VOC to THC[a]	N days with VOC-derived 'THC' TWA estimates	N days with a VOC or THC measurement	Percentage increase of covered days due to VOC-derived estimates[b]
ROV vessels	BOA Deep C	49	14	30	79	61.2
	BOA Sub C	63	15	21	84	33.3
	Casey Chouest	8	2	5	13	62.5
	Chouest Holiday	35	0	27	62	77.1
	Helix Express	10	0	48	58	480
	HOS Achiever	32	0	42	74	131.3
	Iron Horse	11	0	12	23	109.1
	Normand Fortress	17	11	0	17	0
	Ocean Intervention I	15	2	2	17	13.3
	Ocean Intervention III	90	67	25	115	27.8
	Olympic Challenger	39	2	18	57	46.2
	REM Forza	77	9	26	103	33.8
	Skandi Neptune	110	38	16	126	14.5
	All ROV vessels	556	160	272	828	49
Marine vessels	Adriatic	1	0	38	39	3800
	Blue Dolphin	8	0	26	34	325
	Helix Producer	9	0	26	35	288.9
	HOS Strongline	15	0	9	24	60
	Loch Rannoch	16	0	4	20	25
	Massachusetts	10	0	19	29	190
	Monica Ann	40	0	24	64	60
	Normand Commander	54	0	7	61	13
	Odyssea Diamond	28	0	23	51	82.1
	Overseas Cascade	2	0	11	13	550
	Seacor Pride	19	0	4	23	21.1
	Seacor Reliant	12	0	1	13	8.3
	Seacor Vanguard	2	0	9	11	450
	Seacor Venture	35	0	3	38	8.6
	Stim Star 3	10	0	3	13	30
	Tyler Stephen	57	0	11	68	19.3
	War Admiral	20	0	5	25	25
	All marine vessels	338	0	223	561	66.0
All marine and ROV vessels		894	160	495	1389	56.6

[a]This number (*N*) is included in the THC personal measurement day count as it is the number of days with at least three non-censored THC samples with corresponding VOC measurements coverage.

[b][(Total number of days with THC measurements or 'THC' estimates/*N* THC measurements) × 100] − 100 to obtain the increase above the THC measurements.

2 (available at *Annals of Work Exposures and Health* online), including the final THC estimates for each of the ROV vessels and the other MVs with VOC measurements. Document 2 also identifies the number of THC measurements and THC-derived estimates for each time period, the percentage of measurements <LOD, and the AMs, GMs, GSDs, and the 95th percentiles and their credible intervals.

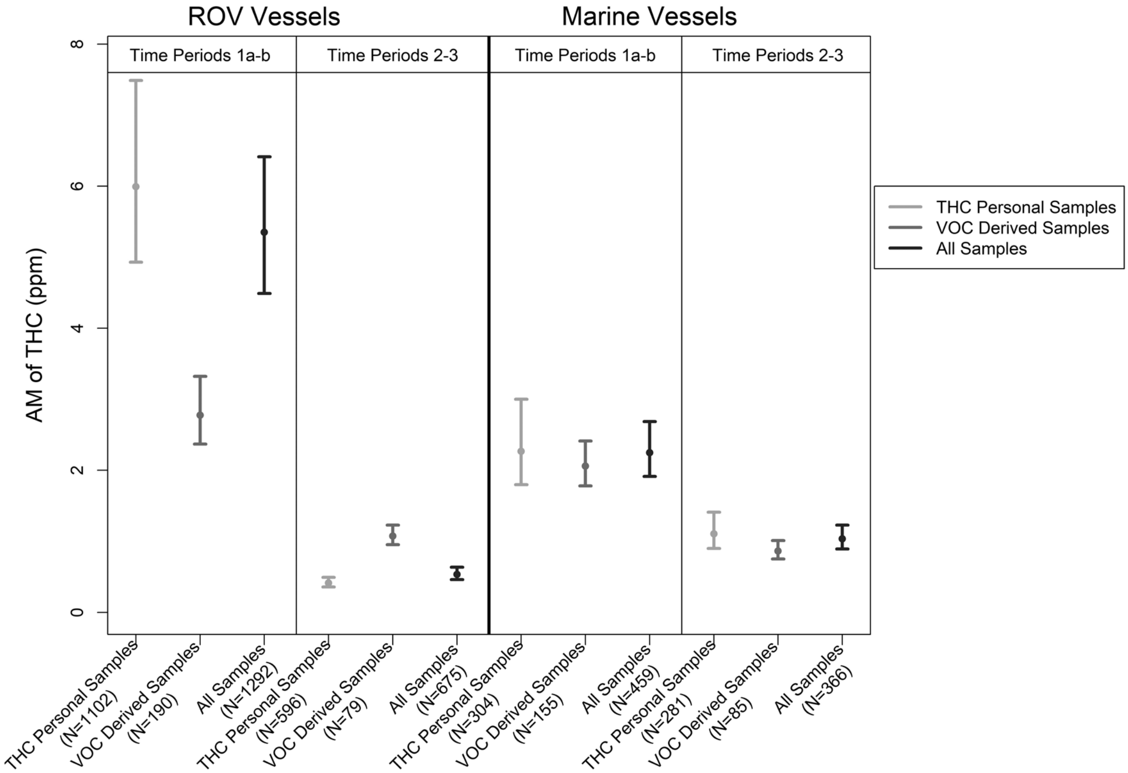


Figure 3. AMs of THC from THC personal samples alone, of the VOC measurements, and of all samples (THC + VOC-derived "THC" TWA estimates). The point represents the median with the bars representing the 95% credible intervals.

Supplementary Table S7, document 1 (available at *Annals of Work Exposures and Health* online), identifies the between and within vessel-day variances for the ROV vessels and MVs for the VOC data. The hour-to-hour variability within a day on a vessel was less than the variation observed between days. The GSD of the within vessel-day variability within the ROV vessels was 2.66, while the GSD of the between vessel-day variability was 3.89. For the MVs, the GSD of within vessel-day variability was 2.27, while the GSD of between vessel-day variability was 2.81.

Discussion

Unlike previous oil spills, the *DWH* oil spill was characterized by a large number of personal exposure measurements. However, the response to the spill was also one of the largest, with rostering of ~55 000 workers (NIOSH, 2011) who performed a large number of activities over a vast expanse of water and land for over 14 months. Area VOC measurements were available from real-time instruments located on several large ROV vessels and MVs in the immediate area of the leaking wellhead in the time periods when the highest exposures likely occurred. These area measurements were used to supplement the personal THC exposure measurements.

There were several challenges to using these real-time measurements. First, it cannot be automatically assumed that area measurements are good measures of personal exposures for a variety of well-established reasons. Second, there was no documentation of the locations where these instruments were placed on each vessel, including if they were indoors or outdoors. The reasons the RP indicated for performing the monitoring suggest that the data were not collected to be representative of full-shift personal exposures but were collected to quickly identify if some locations had high concentrations that could lead to unacceptable personal exposures. This is unlike a situation where an industrial hygienist might use direct-reading instruments to facilitate the control of exposures, e.g. to assess the efficacy of controls in minimizing emissions from an emission source, in which case direct-reading instrumentation may not be appropriate for characterizing worker exposures. Third, we have no other information on the sampling strategy used for the VOC data collection.

Table 3. AM, GM, and GSD estimates for the combined THC measurements and VOC-derived 'THC' TWA estimates for the 13 ROV vessels and 17 MVs over the two sets of time periods in areas other than land. Shaded rows represent vessels with fewer than five observations where no estimates were derived. The percent censored column represents the number of censored THC measurements as a percent of total (THC + VOC) number of measurements.

Vessel type	Vessel	Time periods 1a–1b					Time periods 2–3				
		N	% censored	AM (ppm) median	GM (ppm) median	GSD median	*N*	% censored	AM (ppm) median	GM (ppm) median	GSD median
ROV vessels	BOA Deep C	141	14.9	6.67	1.02	7.24	85	45.9	0.32	0.12	4.11
	BOA Sub C	127	22.0	16.19	1.23	10.81	133	36.8	0.25	0.15	2.87
	Casey Chouest	28	17.9	0.84	0.44	3.25					
	Chouest Holiday	21	14.3	0.75	0.51	2.49	64	21.9	1.01	0.41	3.94
	Helix Express	34	0.0	0.95	0.80	1.82	32	50.0	0.52	0.14	5.47
	HOS Achiever	23	0.0	1.78	1.38	2.08	85	16.5	0.62	0.35	2.88
	Iron Horse	24	25.0	2.60	0.66	5.74	10	0.0	0.85	0.44	3.49
	Normand Fortress	71	1.4	1.71	0.84	3.33	12	16.7	1.26	0.34	6.03
	Ocean Intervention 1	48	54.2	0.28	0.10	4.54	4	50.0	0.13	0.09	2.90
	Ocean Intervention III	314	3.2	5.42	1.74	4.55	76	5.3	0.86	0.53	2.68
	Olympic Challenger	58	37.9	0.97	0.27	5.16	35	37.1	1.05	0.24	6.03
	REM Forza	110	16.4	6.76	1.00	7.44	65	81.5	0.24	0.03	8.09
	Skandi Neptune	293	19.1	7.20	0.83	8.41	74	43.2	0.39	0.13	4.47
	All	1292	15.2	5.35	0.94	6.57	675	35.3	0.54	0.20	4.13
Marine vessels	Adriatic	16	0.0	0.58	0.55	1.41	19	0.0	0.40	0.40	1.01
	Blue Dolphin	30	0.0	2.69	1.84	2.44	16	0.0	0.92	0.85	1.51
	Helix Producer	25	24.0	2.93	0.66	6.21	19	31.6	0.34	0.18	3.21
	HOS Strongline	15	6.7	13.40	2.92	6.74	16	6.3	0.43	0.38	1.64
	Loch Rannoch	29	13.8	1.57	0.59	4.23	8	0.0	1.65	1.22	2.30
	Massachusetts	16	6.3	3.68	0.89	6.14	8	0.0	1.20	0.60	3.77
	Monica Ann	45	6.7	3.12	1.30	3.86	48	35.4	0.89	0.22	5.68
	Normand Commander	42	7.1	1.49	0.64	3.75	56	12.5	1.33	0.47	4.32
	Odyssea Diamond	19	5.3	1.81	1.15	2.68	39	0.0	1.07	0.68	2.61
	Overseas Cascade	14	7.1	2.32	0.99	4.01					
	Seacor Pride	19	5.3	0.54	0.41	2.19	17	5.9	1.48	0.55	4.41
	Seacor Reliant	1	0.0								
	Seacor Vanguard	16	0.0	5.65	3.62	2.64					
	Seacor Venture	31	0.0	3.23	1.42	3.71	32	3.1	1.63	0.88	3.12
	Stim Star 3	28	28.6	0.75	0.32	3.80	14	0.0	1.77	1.54	1.71
	Tyler Stephen	76	5.3	1.98	0.87	3.67	46	10.9	0.87	0.58	2.52
	War Admiral	2	0.0				9	0.0	1.52	1.18	2.13
	All	459	7.2	2.25	0.95	3.73	366	12.3	1.03	0.51	3.28
ROV and marine vessels (all)		1751	13.1%	4.21	0.95	5.69	1041	27.2%	0.73	0.28	4.02

Fourth, while extensive area measurements may be easier to collect than personal exposures, these data are much more difficult to summarize and relate to (particularly full-shift) personal exposure levels. As a result, very few statistical strategies have been proposed to transform these often large datasets of short-duration concentrations into meaningful exposure statistics. Fifth, the start and stop times were non-uniform from one day to the next nor were the sampling rates of the instruments uniform. Nevertheless, even under these limited conditions, we demonstrated empirically that there was a strong Pearson correlation (ρ) between the two types of measurements (ρ = 0.78), such that the VOC measurements could be used to estimate personal exposures.

We considered various options for dealing with the lack of information on the instrument location, such as weighting instrument results by the number of observations. This approach, however, carried the (possibly unlikely) risk that an instrument with many observations was in a location with minimal exposure, while another instrument with fewer observations was located where exposure levels were higher. Considering that it was important to know that living, eating, recreation, and office areas were virtually free of oil-related vapors, the RP guidance specifically mention that some of the instruments were to be located in these areas. We also considered removing any instruments on any day when the values were all below the LOD, on the assumption that the instrument was located indoors. This approach, however, could have resulted in removing measurements that actually reflected outdoor locations with low air concentrations, e.g. under high wind conditions. Also, the workers wore the passive dosimeters for the entire day including in break rooms or while eating in the ships galley which were inside areas. Given the lack of information, we chose to weight each instrument equally in calculating this average.

A key factor in being able to supplement the THC personal samples where there were gaps in measured days was the relationship between the (minimum 4-h) VOC TWAs and the full-shift personal THC measurements. Groth, Banerjee *et al.* (2021) described the procedure to convert the real-time VOC collected at different sampling hours and at different sampling rates to form a set of 21 900 hourly averages, which were then used to develop daily TWAs to match to the THC sampling periods. We required at least three personal full-shift THC measurements with corresponding paired VOC daily averages for a vessel on a single day to consider the day in the VOC:THC comparison and at least 3 days in the time period for the regression analysis. These criteria increased the likelihood that our estimates represented workers' exposures over each entire vessel. We chose time periods 1a and 1b combined (i.e. prior to 16 July 2010, when the oil was being released) when both types of measurements were most often available and obtained a regression equation between them. The reasonably strong and significant relationship between them for this time period (slope = 0.63 (95% CI 0.54, 0.71), $\rho = 0.78$) and large sample size ($n = 131$ vessel-days) gave us confidence that using it to predict estimates was reasonable. Support for applying this relationship to estimate exposures in other time periods and vessels includes (i) the same general sampling strategy likely used for the direct-reading instruments (even though we do not know the actual strategy) and (ii) the locations that a THC measured worker spent for a portion of a workday were probably the same as the locations measured for VOC. Moreover, unlike workers on the rig vessels (Huynh *et al.*, 2021b), these exposure estimates were assigned to all study participants on the vessel and not at the level of the job title, because job title information was typically not available for these GuLF STUDY participants. Thus, an average of stationary instruments situated throughout the vessel to reflect, at least to some degree, where workers were located across a vessel was likely to be at least somewhat representative of all workers on the vessel.

The procedure substantially increased the number of days for which daily exposure estimates were available. It is important to note that we only had general information on the deployment dates of the vessels, so we cannot estimate the fraction of the days of operation for which we have estimates. The number of vessel-days for which we have exposure estimates, nonetheless, increased from 556 to 828 (a 49% increase) for ROV vessels and from 338 to 561 (a ~66% increase) for MVs and more importantly, from 414 to 606 for the ROV vessels in TP1a–1b, on which we based our prediction. If our assumption that the observed correlation between VOC and THC holds true for the days with no THC, this increase in sample size should provide us with more representative estimates of the distribution of the measured exposure levels, and thus, a better estimate of the AM than without the VOC data. The correlation assumption is based on the chemical and physical properties of the crude oil and the oil's degree of weathering. The time period accounts for weathering. Therefore, the day-to-day correlation between VOC and THC should hold no matter if a THC measurement was collected or not.

Fig. 3 shows that the overall AM exposures on ROV vessels and MVs significantly declined from TP1a–1b to TP2–3. This decline in concentrations over time is consistent with similar declines observed in the rig vessels (Huynh *et al.*, 2021b) and with other vessels (Huynh *et al.*, 2021a). The likely cause for the high levels in TP1a–1b is the large amount of uncontrolled oil coming from the well [estimated to be ~50 000 barrels/day (McNutt *et al.*, 2011)] during this period. Much of the released oil rose to the water surface and evaporated, leading to relatively substantial vapor concentrations (Huynh *et al.*, 2021a,b). The well was successfully mechanically capped on 15 July (the end of TP1b), and from that time on, little new oil was released from the BOP, leading to the substantial decrease in concentrations of vapors in TP2–3 and later.

We observed credible differences between the AMs of the ROV vessels and the MVs in TP1a–1b. Because AMs are a function of both GM and GSD, the difference was more likely due to the greater variability in exposures in the ROV vessels (overall GSD in TP1a–1b = 6.57) compared to MVs in the same time period (GSD = 3.73). This greater variability in the exposures on the ROV vessels is likely due to a wide variety of tasks that these vessels were undertaking during TP1a–1b, resulting in a greater contrast of exposures among the workers on the ROV vessels, i.e. some workers having (relatively) higher exposures while cleaning the ROVs and other workers having lower exposures due to their operating the ship. In contrast, on the MVs the difference of exposures may have been smaller due to fewer high exposed tasks, so that their main source of exposure was the ambient air, resulting in more homogeneous exposures across the MV workers. The GSD estimates, particularly in TP1a–1b for the ROV vessels, were greater than typical values observed in most stable occupational scenarios. This is likely due to the non-routine nature of the oil spill clean-up tasks from one day to the next, especially in TP1a–1b, and to the fact that while most of the work would have been outside, some of the measurements may have been indoors.

The new THC AM exposure estimates on the ROV vessels were generally higher compared to the MVs, especially so in TP1a–1b, but the difference between the two types of vessels is much smaller in TP2–3. The overall GMs for ROV vessels and MVs during TP2–3 are not substantially different (but remain credibly different; 0.20 and 0.51 ppm, respectively). The cessation of the oil release beginning TP2 likely lessened the ambient air levels and the variability of tasks, resulting in lower GSD values for both ROV vessels and MVs (although still on the high side compared to more typical occupational settings). We have found similarly high GSDs for other workers in the study (Huynh *et al.*, 2021a,b,c).

The reasons for the differences between the vessels shown in Fig. 3 may be several but may lie with the MVs. A THC:VOC correlation comparable to that found with the ROV vessels was found on two of the four rig vessels, *Discoverer Enterprise* and *Development Driller III*. On these two vessels, we found virtually an identical correlation in TP1a–1b as that with the ROV vessels ($\rho = 0.73$), suggesting that the correlation was appropriate for vessels in the wellhead area (Supplementary Table S8, document 1, available at *Annals of Work Exposures and Health* online). We did not use the VOC data on these two rig vessels to estimate days with no THC measurements, however, because we had sufficient measurements on those vessels to generally estimate exposures by job groups. Also it was unclear how to assign area measurements to specific jobs (in contrast to the ROV and MV assignments at the vessel level). In addition, the rig correlation served as a useful reference value to our observed ROV vessel correlation.

For the MVs in Fig. 3, the difference between the AMs of the THC personal samples and the AMs incorporating the ‘THC’ TWA estimates that was seen between TP1a–1b and TP2–3 for the ROV data was not observed. The failure to observe differences in the MV AMs may be that it was inappropriate to apply the ROV vessels correlation. As indicated above, the similar intercepts and slopes found for the two rig vessels (Supplementary Table S8, document 1, available at *Annals of Work Exposures and Health* online), however, suggest that this is not the reason. One possible reason is that the MVs had fewer observations to evaluate, which may have resulted in a weaker correlation. It may be that the function of the MVs and the activities performed by workers resulted in workers not coming into as close contact with the oil as the ROV vessel workers, resulting in less of a contrast between high and low exposed workers on the vessel (as exhibited by the lower GSD on the MVs). This situation would have reduced our ability to see a difference between the two sets of time periods. The ROV vessel correlations, therefore, were used to predict THC personal exposures for the MVs, but as a result, we have less confidence in our estimates for the MVs.

To provide some sense of internal validity, we evaluated the change in the AMs based on the proportion of added measured vessel-days due to the VOC ‘THC’ TWA estimates. We arbitrarily identified as >65% [determined by dividing the number of days with THC personal measurements by the total number of measured (THC+VOC-derived ‘THC’ TWAs) days], as the cutoff to indicate a small increase in THC coverage. That is, if the original THC measurements contributed >65% of the final number of observations, we did not expect the AMs to change substantially after adding the VOC ‘THC’ estimates. In contrast, <65% coverage indicated a large increase in the number of measurements over the original THC measurements, which was likely to have a greater impact on the AM. For vessels where the cutoff was >65%, there was little change in the AMs. This suggests that the methodology we used here was appropriate and that the VOC measurements were representative of the THC exposures. For those vessels where the coverage of days monitored was lower (<65%), the inclusion of the VOC data more often changed the value, likely improving the final THC estimates. This overall difference is shown in the ROV vessel data in Fig. 3, where, for the two periods evaluated (1a–1b and 2–3), the THC exposure estimates from the THC personal measurements alone were observed to be different from final observations (although significance cannot be judged because the groups are not independent).

To develop the daily exposure estimates of VOC, we averaged hourly estimates on each day. Results from our ANOVA revealed that the hour-to-hour variability on each vessel-day was less than the variability between daily averages on different vessels. This result suggests that while we eliminated variability in VOC concentrations by developing hourly averages, this variation would not have increased the overall variability of the estimates substantially. Moreover, the variability in the VOC measurements that we lost is not likely to differ from the variability lost in a full-shift TWA measurement. Finally, the impact of not incorporating hour-to-hour variability is likely to be minimal since we had no information on what study participants were doing on any vessel (or where), resulting in all participants being assigned the same value over a time period. Our methods, however, account for the larger source of day-to-day variability on different vessels and the vessel-to-vessel variability.

There are currently no OELs for THC. However, petroleum distillates, which are the closest equivalent, has a NIOSH recommended exposure limit of 86 ppm (or 350 mg m^{-3}) (National Institute of Occupational Safety and Health, 2007). The THC estimates obtained for the ROV vessels and MVs were substantially below these exposure limits.

As part of this work, we used a Bayesian linear regression framework to develop additional THC observations. This model assumes lognormality of the THC and VOC measurements, equal variances, linearity, and independence of observations. Non-linear relationships or other distributional assumptions were not considered. Autocorrelation between measurements at consecutive time points is commonly seen when analyzing time-series data and may need to be adjusted for when comparing exposure estimates. The analyses described in Groth, Banerjee *et al.* (2021) do not account for autocorrelation that may be present between one direct-reading observation to the observation directly following it. While it is possible that some of the VOC estimates (described in Groth, Banerjee *et al.*, 2021) may have been influenced by autocorrelation, the primary purpose of Groth, Banerjee *et al.* (2021) and this paper is to use the time-series VOC measurements to estimate THC exposure, not to characterize VOC exposures. In the estimation of study participants' exposures on these vessels, there are several sources of uncertainty in addition to possible autocorrelation. Each instrument's real-time measurements is used to create 1-h VOC averages; then the hourly averages from multiple instruments at multiple locations are averaged over the hours sampled by their paired personal THC measurements. The concentrations at these various locations are likely not correlated. The averaging over multiple instruments (locations) as well as multiple hours and days will mitigate the effect of fine temporal resolution autocorrelation. While there can be potential advantages for an autocorrelation model (temporal and perhaps even spatial), the scope for such an analysis is beyond the objectives of this particular manuscript. Future methods could consider autocorrelation, as in O'Shaughnessy and Cavanaugh (2015), in the development of the hourly averages.

The THC AMs for these vessels are some of the highest observed in the GuLF STUDY, and the levels were of the same order or greater than some of those observed on the rig vessels (Huynh *et al.*, 2021b). Despite these relatively (to the study) substantial exposures, personal sampling on these vessels was not conducted to the same extent as on the rig vessels, making it even more important that exposures for this group of workers be assessed using the additional measurements from the VOC-reading instruments. But as shown in Fig. 3, the median estimate across all vessels decreased as a result of our using the VOC data, suggesting that at least for some of the vessels workers' exposures would have been overestimated had we not done this work. The potential for such relatively high exposures for the study participants on these vessels could have an impact on the health risk outcomes investigated in the epidemiologic study. Lower exposed workers misclassified to high exposures levels can result in attenuating a disease risk to the null.

This paper presents a novel method that facilitated the use of real-time area measurements to supplement personal exposure measurements for developing (in our case) likely representative estimates of some of the more highly exposed workers aboard some of the vessels involved in the OSRC activities. This method was feasible because we found a good correlation between the daily averages of the VOC-derived TWAs and the THC measurements. The availability of such high-volume real-time measurement data occurs with some frequency in occupational settings where real-time, as well as personal, measurements for vapors, gases, and aerosols are collected. There is an understandable reluctance to use these real-time measurements for exposure assessment. First, if real-time measurements are reflecting the efficacy of engineering controls, they may not be appropriate for personal exposures. Second, they can generate very large databases of measurements that are not easily analyzable. In fact, we had to use the University of Minnesota Supercomputing Institute and parallel computing to develop these estimates. Also, area static samplers can under- or overestimate personal exposures. Nonetheless, the use of real-time monitoring is likely to increase significantly in the near future with rapid improvements in low-cost sensors and monitoring devices and developments in handling high-volume data. The methods described in this paper and Groth, Banerjee *et al.* (2021) could be used to manage real-time measurements for exposure assessment purposes, especially in cases where a correlation might be expected between area and personal measurements. Groth, Huynh *et al.* (2021) extended the use of the correlations by calculating correlations between the new 'THC' TWA observations and each of the BTEX-H substances to estimate the latter for the same vessel-days as we did for THC.

Conclusions

We used the database of hourly real-time VOC averages described by Groth, Banerjee *et al.* (2021) to develop daily vessel averages for use in the estimation of THC exposures for workers on two types of vessels—ROV vessels and MVs—in the GuLF STUDY. This was possible, in large part, due to the amount of VOC and THC measurements collected on the same vessels and the strength of the correlation between the full-shift THC and the VOC-derived 'THC' TWA daily averages. These 'THC' TWA estimates substantially supplemented the THC personal measurements, with the number of vessel-days for which we have daily exposure estimates, increasing by ~60%, thereby increasing the likelihood of more representative estimates AM and the distribution of exposures that should better reflect the working conditions of the workers on these vessels. These AM estimates are some of the highest observed in the GuLF STUDY, with measurements comparable to or higher than those found on some of the rig vessels, so it was crucial to have confidence that these high levels were reflective of true exposures for the epidemiologic analysis. The AM estimates decreased over time, consistent with findings for other workers in our study, likely caused by the capping of the well that reduced oil vapors. The method used here can be applied to other occupational instances where large amounts of area measurements can be used for personal exposure estimation. The AMs generated by this process have been used in job-exposure matrices in the GuLF STUDY.

Supplementary Data

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

Funding

Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01 ES102945). Dr Sudipto Banerjee was also supported by the following NIH and NSF grants during this work: NIH/NIEHS 1R01ES027027-01; NIH/NIEHS R01ES030210-01; and NSF DMS-1513654, NSF IIS-1562303, and NSF DMS-1916349. Dr Tran Huynh was also supported, in part, by CDC/NIOSH award K01OH011191 award.

Acknowledgements

We thank BP for providing the measurements and Wendy McDonald and Caitlin Rousch of McDowell Safety and Health Services, Inc. for their tremendous efforts in organizing the data. We would like to acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. We also thank BP for providing us with these data for analysis

Conflict of interest

The authors declare no conflict of interest relating to the material presented in this article. Its contents, including any opinions and/or conclusions expressed, are solely those of the authors.

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.

BP Gulf Science Data. (2016) *Monitoring of the Personal Breathing Zone of Response workers for chemicals and oil mists from April 2010 to January 2012. Distributed by Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University–Corpus Christi. BP Science, SRN-2001.1-WorkPlan.pdf. Data.* doi:10.7266/N71G0JQK. UDI: BP.x750.000:0007

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 C, Banerjee S, Ramachandran G *et al*. (2017) Bivariate left-censored Bayesian model for predicting exposure: preliminary analysis of worker exposure during the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; **61**: 76–86.

Groth CP, Banerjee S, Ramachandran G *et al*. (2018) Multivariate left-censored Bayesian modeling for predicting exposure using multiple chemical predictors. *Environmetrics*; **29**: 1–16.

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 inhalation exposures to oil-related components on the supporting vessels during the *Deepwater Horizon* oil spill. *Ann Work Expo Health*; **66**: i111–123.

Huynh TB, Groth CP, Ramachandran G *et al*. (2021b) 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*. (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*. (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.

McNutt M, Camilli R, Guthrie G *et al*. (2011) *Assessment of flow rate estimates for the Deepwater Horizon/Macondo Well oil spill. Flow Rate Technical Group Report to the National Incident Command, Interagency Solutions Group, 10 March 2011*.

National Institute of Occupational Safety and Health. (2007) *NIOSH Guide to Chemicals*. DHHS (NIOSH) Publication No. 2005-149. Available at https://www.cdc.gov/niosh/npg/default.html. Accessed 16 June 2021.

NIOSH. (2011) *NIOSH Deepwater Horizon Roster Summary Report*. National Institute for Occupational Safety and Health. Available at https://www.cdc.gov/niosh/updates/upd-12-19-11.html. Accessed 16 June 2018.

O'Shaughnessy P, Cavanaugh JE. (2015) Performing *T*-tests to compare autocorrelated time series data collected from direct-reading instruments. *J Occup Environ Hyg*; **12**: 743–52.

Plummer M. (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), 20–22 March, Vienna, Austria.

Plummer M. (2016) rjags: Bayesian graphical models using MCMC. R package version 4-6. Available at https://cran.r-project.org/web/packages/rjags/index.html. Accessed 12 June 2021.

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.

R Core Team. (2017) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/ Accessed 12 July 2021.

Stenzel MR, Arnold SF, Ramachandran G *et al*. (2021) Estimation of airborne vapor 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, Banerjee S *et al*. (2021) Exposure assessment techniques applied to the highly censored *Deepwater Horizon* Gulf oil spill personal measurements. *Ann Work Expo Health*; **66**: i56–i70.

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 PA, Gorman Ng M, Cherrie JW *et al*. (2021) Estimation of dermal exposure to oil spill response and clean-up workers after the *Deepwater Horizon* disaster. *Ann Work Expo Health*; **66**: i234–i246.

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