Conservation Biology

CONTRIBUTED PAPER

Modeling population effects of the *Deepwater Horizon* oil spill on a long-lived species

Lori H. Schwacke[1] | **Tiago A. Marques**[2] | **Len Thomas**[2] | **Cormac G. Booth**[3] |
Brian C. Balmer[1,#] | **Ashley Barratclough**[4] | **Kathleen Colegrove**[5] | **Sylvain De Guise**[6] |
Lance P. Garrison[7] | **Forrest M. Gomez**[4] | **Jeanine S. Morey**[1] | **Keith D. Mullin**[8] |
Brian M. Quigley[1] | **Patricia E. Rosel**[9] | **Teresa K. Rowles**[10] | **Ryan Takeshita**[1] |
Forrest I. Townsend[11] | **Todd R. Speakman**[1] | **Randall S. Wells**[12] | **Eric S. Zolman**[1] |
Cynthia R. Smith[4]

[1] National Marine Mammal Foundation, Johns Island, South Carolina, USA

[2] Centre for Research into Ecological and Environmental Modelling (CREEM), University of St Andrews, The Observatory, London, UK

[3] SMRU Consulting, Scottish Oceans Institute, East Sands, University of St Andrews, St Andrews, UK

[4] National Marine Mammal Foundation, San Diego, California, USA

[5] Zoological Pathology Program, College of Veterinary Medicine, University of Illinois at Urbana-Champaign, Brookfield, Illinois, USA

[6] Department of Pathobiology and Veterinary Science, University of Connecticut, Storrs, Connecticut, USA

[7] National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, Miami, Florida, USA

[8] National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, Pascagoula, Mississippi, USA

[9] National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, Lafayette, Louisiana, USA

[10] National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Office of Protected Resources, Silver Spring, Maryland, USA

[11] Bayside Hospital for Animals, Fort Walton Beach, Florida, USA

[12] Chicago Zoological Society's Sarasota Dolphin Research Program, c/o Mote Marine Laboratory, Sarasota, Florida, USA

Correspondence
Lori H. Schwacke, National Marine Mammal Foundation, 3419 Maybank Highway, Suite B, Johns Island, SC 29455, USA.
Email: lori.schwacke@nmmf.org

\#Present address: Dolphin Relief and Research, 6 Antelope Way, Clancy, Montana 59634, USA.

Article impact statement: Oil spills can have long-term consequences for the health of long-lived species; thus, effective restoration and monitoring are needed.

Abstract

The 2010 *Deepwater Horizon* (DWH) oil spill exposed common bottlenose dolphins (*Tursiops truncatus*) in Barataria Bay, Louisiana to heavy oiling that caused increased mortality and chronic disease and impaired reproduction in surviving dolphins. We conducted photographic surveys and veterinary assessments in the decade following the spill. We assigned a prognostic score (good, fair, guarded, poor, or grave) for each dolphin to provide a single integrated indicator of overall health, and we examined temporal trends in prognostic scores. We used expert elicitation to quantify the implications of trends for the proportion of the dolphins that would recover within their lifetime. We integrated expert elicitation, along with other new information, in a population dynamics model to predict the effects of observed health trends on demography. We compared the resulting population trajectory with that predicted under baseline (no spill) conditions. Disease conditions persisted and have recently worsened in dolphins that were presumably exposed to DWH oil: 78% of those assessed in 2018 had a guarded, poor, or grave prognosis. Dolphins born after the spill were in better health. We estimated that the population declined by 45% (95% CI 14–74) relative to baseline and will take 35 years (95% CI 18–67) to recover to 95% of baseline numbers. The sum of annual differences between baseline and injured population sizes (i.e., the lost cetacean years) was 30,993 (95% CI 6607–94,148). The population is currently at a minimum point in its recovery trajectory and is vulnerable to emerging threats, including planned ecosystem restoration efforts that are likely to be detrimental to

the dolphins' survival. Our modeling framework demonstrates an approach for integrating different sources and types of data, highlights the utility of expert elicitation for indeterminable input parameters, and emphasizes the importance of considering and monitoring long-term health of long-lived species subject to environmental disasters.

Article impact statement: Oil spills can have long-term consequences for the health of long-lived species; thus, effective restoration and monitoring are needed.

KEYWORDS

dolphin, expert elicitation, health assessment, marine mammal, oil spill, population model, slow-living species

Resumen

El derrame de petróleo *Deepwater Horizon* (DWH) en 2010 expuso gravemente a este hidrocarburo a los delfines (*Tursiops truncatus*) de la Bahía Barataria, Luisiana, causando un incremento en la mortalidad y en las enfermedades crónicas, y deteriorando la reproducción de los delfines sobrevivientes. Realizamos censos fotográficos y evaluaciones veterinarias durante la década posterior al derrame. Asignamos un puntaje pronóstico (bueno, favorable, moderado, malo, o grave) a cada delfín para proporcionar un indicador integrado único de la salud en general. También examinamos las tendencias temporales de estos puntajes. Usamos información de expertos para cuantificar las implicaciones de las tendencias para la proporción de delfines que se recuperaría dentro de su periodo de vida. Integramos esta información, junto con información nueva, a un modelo de dinámica poblacional para predecir los efectos sobre la demografía de las tendencias observadas en la salud. Comparamos la trayectoria poblacional resultante con aquella pronosticada bajo condiciones de línea base (sin derrame). Las condiciones de enfermedad persistieron y recientemente han empeorado en los delfines que supuestamente estuvieron expuestos al petróleo de DWH: 78% de aquellos evaluados en 2018 tuvieron un pronóstico moderado, malo o grave. Los delfines que nacieron después del derrame contaron con mejor salud. Estimamos que la población declinó en un 45% (95% CI 14–74) relativo a la línea base y tardará 35 años (95% CI 18–67) en recuperar el 95% de los números de línea base. La suma de las diferencias anuales entre el tamaño poblacional de línea base y el dañado (es decir, los años cetáceos perdidos) fue de 30,993 (95% CI 6,607-94,148). La población actualmente está en un punto mínimo de su trayectoria de recuperación y es vulnerable a las amenazas emergentes, incluyendo los esfuerzos de restauración ambiental planeada que probablemente sean nocivos para la supervivencia de los delfines. Nuestro marco de modelado demuestra una estrategia para la integración de diferentes fuentes y tipos de datos, resalta la utilidad de la información de expertos para los parámetros de aportación indeterminable, y enfatiza la importancia de la consideración y el monitoreo de la salud a largo plazo de las especies longevas sujetas a los desastres ambientales.

Modelado de los Efectos Poblacionales del Derrame de Petróleo *Deepwater Horizon* sobre Especies Longevas

PALABRAS CLAVE

análisis de salud, delfín, derrame de petróleo, especie longeva, información de expertos, mamífero marino, modelo poblacional

INTRODUCTION

In 2010, an explosion on the *Deepwater Horizon* (DWH) oil platform killed 11 people and led to the largest marine oil spill in U.S. history. Approximately 130 million gallons of oil flowed into the Gulf of Mexico, spreading to over 2000 km of coastline (Westerholm & Rauch, 2016). Some impacts on wildlife were immediately obvious; thousands of turtles and marine birds were found floating in oil slicks or washed ashore mired in oil.

Other, less direct but more enduring, impacts emerged later. Strandings of dead dolphins with pathological findings consistent with oil-related injury increased (Venn-Watson et al., 2015) and remained elevated for at least 3 years, contributing to the largest cetacean unusual mortality event ever recorded for the Gulf of Mexico (Litz et al., 2014).

In the years immediately following the spill (2010–2015), a natural resource damage assessment (NRDA) was conducted, which included studies to assess injuries to cetacean

populations and identify actions to restore injured resources (Westerholm & Rauch, 2016). The cetacean NRDA field studies focused on common bottlenose dolphins (*Tursiops truncatus*) (hereafter dolphins) because of their accessibility in nearshore waters. Health assessment studies on dolphins in Barataria Bay (BB), Louisiana, where heavy oiling occurred, documented sublethal health effects, including reproductive failure, lung disease, and impaired stress response (Kellar et al., 2017; Schwacke et al., 2014).

As part of the NRDA, a population model was developed to evaluate the time it would take for the BB dolphin population to recover (Schwacke et al., 2017). The model integrated estimates of immediate postspill mortality and reproductive failure derived from photographic identification (photo-ID) studies conducted over 4 years following the spill (Kellar et al., 2017; McDonald et al., 2017) with veterinarians' predictions of how long reproductive impairment and compromised health, which decreased survival, would likely persist. The model was based on the assumption that fecundity rates are density dependent (DD), increasing as a function of reduced population size, and that dolphins born after the spill do not exhibit oil-related health effects. These assumptions were important because recovery of the population is highly dependent on emergence of a new, healthy cohort. Yet, potential for multigenerational health effects was unknown, and the largest sources for uncertainty in population projections were associated with the DD function and estimates of how long it would take for exposed dolphins to return to baseline health status.

Additional years (2016–2019) of photo-ID and health assessment studies now provide a decade of data to assess temporal trends in health and to evaluate population model assumptions regarding multigenerational effects and individual health recovery. We used these additional data and results from formal expert elicitation (EE) to refine population model parameters for recovery to baseline health status and DD fecundity. Using the refined model, we conduct analyses to reexamine the predicted population trajectory, with an aim to provide a better informed assessment for recovery of the BB population. This is timely because ecosystem restoration projects funded by the DWH NRDA settlement are being initiated, which may benefit some natural resources but could have further negative effects on dolphins. Our approach, integrating EE for indeterminable inputs, is uniquely suited for future efforts to assess damage and long-term consequences for long-lived species following oil spills or other environmental disasters, particularly in cases where deaths and births are difficult to observe and for which detection of population trends is difficult.

METHODS

Study areas and data

We used data from studies conducted in BB, including dolphin sighting histories from photo-ID surveys and health diagnostics from capture-release studies. We previously reported results from BB health assessments conducted in 2011, 2013, and 2014 (Schwacke et al., 2014; Smith et al., 2017). Here, we integrated those data with new health assessment data obtained in 2016–2018. Health data were also obtained from capture-release studies at an unoiled comparison site, Sarasota Bay (SB), Florida, in 2011 and 2013–2018. The health data from SB in 2011 and 2013 were previously reported (Schwacke et al., 2014; Smith et al., 2017). For all years, health diagnostics included physical examination, length:mass ratio, serum biochemistry, hematology, and ultrasonography (details in Appendix S1).

Fieldwork was conducted under NMFS permit numbers 18786-03 and 18986-04 (BB) and 522–1785 and 15543 (SB). Protocols were approved by the National Oceanic and Atmospheric Administration and Mote Marine Laboratory Institutional Animal Care and Use Committees for BB and SB, respectively.

Data are publicly available through GOMRI Information & Data Cooperative (GRIIDC) https://data.gulfresearchinitiative.org (https://doi.org/10.7266/N7H41PTV,10.7266/n7-76aj-rp39,10.7266/n7-sv57-1h12).

Temporal trends in health

We compared each dolphin's diagnostic results with established reference intervals to identify health abnormalities, and then experienced veterinarians (C.S., F.G., and F.T.) assigned a prognostic score, providing a single integrated indicator of overall health. As previously defined (Smith et al., 2017), the prognosis categories were good (favorable outcome expected); fair (favorable outcome likely); guarded (outcome uncertain); poor (unfavorable outcome likely); and grave (death considered imminent). Some physical examination abnormalities are within expected limits for wild dolphins—such as minor rake marks, mild tooth wear, or well-healed entanglement injuries. Findings within expected limits from which the animals would survive or make a full recovery were classified as a good or fair prognosis. A guarded prognosis was assigned when the physical examination findings were of concern, such as moderate lung disease, low body mass index, or multiple blood analyte abnormalities. A poor or grave prognosis was assigned when indicators, such as severe lung disease, active infection or serious injuries, and severe blood analyte abnormalities, were present (Smith et al., 2017) and the animal was considered unlikely to survive.

We created an additional binary variable to indicate good or fair versus guarded, poor, or grave. Methods for prognosis scoring are described by Smith et al. (2017) and additional details are in Appendix S1.

Photo-ID images and associated data were analyzed as described by McDonald et al. (2017). We categorized each dolphin as being alive in 2010 during the spill (pre-2010 cohorts) or born after (post-2010 cohorts) (Appendix S1). We hypothesized that BB pre-2010 dolphins have a higher prevalence of guarded-or-worse prognoses than SB pre-2010 dolphins, but that post-2010 dolphins do not differ. We also examined whether BB 2010 dolphin prognosis scores improved over time. For pre-2010 cohorts, we modeled temporal trends in binary prognostic score with a generalized additive model (GAM) (binomial
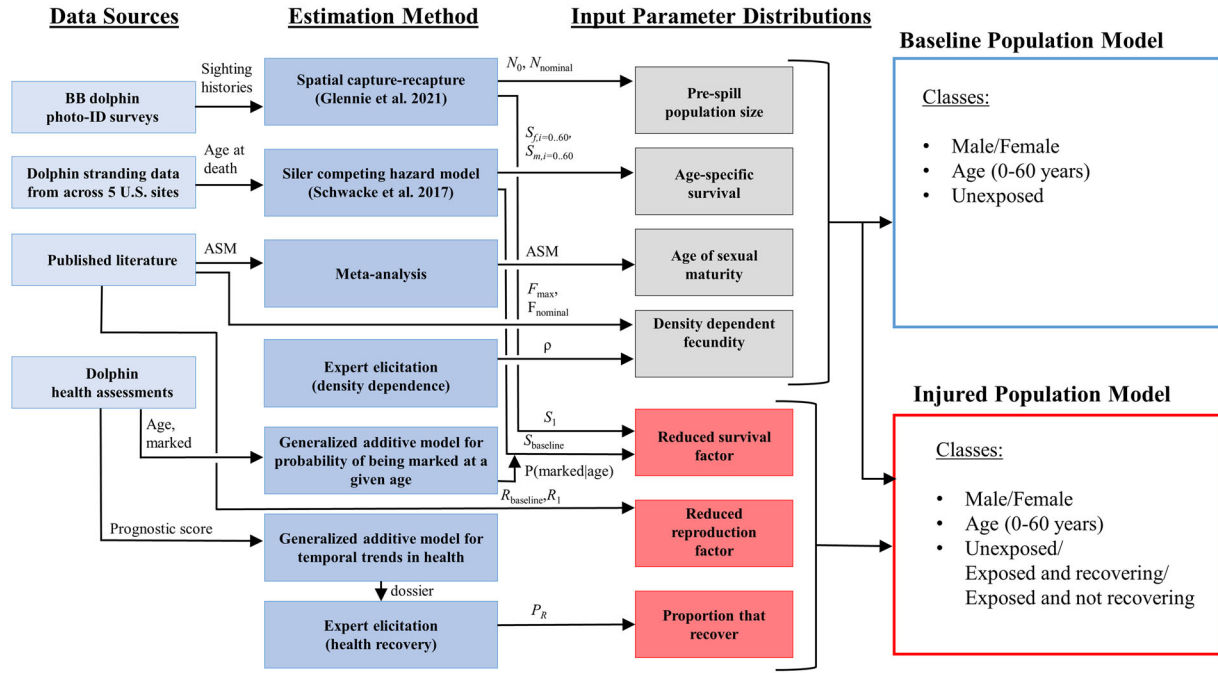
FIGURE 1 Overview and information flow for population model of bottlenose dolphins in Barataria Bay under baseline and injured scenarios

response and logit link) with sex and site (BB or SB) as factor covariates and a separate smooth function of year (2011–2018) for each site (Appendix S1). For post-2010 cohorts, we could not model temporal trends due to limited sample sizes, so we fitted a binomial generalized linear model with covariates sex and site.

Population dynamics model structure

We expanded on the age-, sex-, and class-structured matrix population model described by Schwacke et al. (2017) (Figure 1). We used the same 61 age classes, but rather than the dichotomous exposed or unexposed classes used by Schwacke et al. (2017), we implemented three animal classes. Unexposed animals included dolphins that were not exposed to DWH oil because they were born after the spill and dolphins that were exposed but recovered back to baseline survival and fecundity levels. Exposed and recovering animals had reduced levels of survival and fecundity immediately following the spill (in 2010) but gradually recovered to baseline survival and fecundity levels. Exposed and not recovering animals had reduced levels of survival and fecundity in 2010 and remained at these reduced levels throughout their life.

For unexposed dolphins, the transition from one age to the next was age and sex specific. Survival probabilities were $S_{f,i}$ and $S_{m,i}$ for age $i = 0, \dots, 60$ females and males, respectively. The same survival probabilities were applied for the exposed classes but were adjusted by multiplying by a survival factor (SF), which could change over time. The survival rates for the exposed classes for age i and year t were then:

$$S'_{f,i}(t) = \mathrm{SF}(t) \cdot S_{f,i}$$

and

$$S'_{m,i}(t) = \mathrm{SF}(t) \cdot S_{m,i} \tag{1}$$

for females and males, respectively. For the exposed and not recovering class, SF remained constant over their lifetimes ($\mathrm{SF}(t) = \mathrm{SF}_1$ for any t). For dolphins that eventually recovered (exposed recovering class), SF increased as a function of t:

$$\mathrm{SF}(t) = \min\left[1, \mathrm{SF}_1 + \frac{t-1}{T_R}(1 - \mathrm{SF}_1)\right], \tag{2}$$

where T_R is the number of years before full recovery. Estimations of SF_1 and T_R are detailed below.

Total population size in year t, summed over age, sex, and exposure classes, is denoted as N_t, and the model starts in year $t = 0$, which corresponds to 2010, just before the DWH oil spill. Females reproduce and contribute new members to the 0 age class through a DD fecundity function, $f(N_t)$, which decreases as population size increases and was applied for females once they were above the age of sexual maturity. Similar to survival probability, a probability of successful reproduction was adjusted by a reproductive factor (RF):

$$f'(N_t) = \mathrm{RF}(t) \cdot f(N_t), \tag{3}$$

where $\mathrm{RF}(t) = \mathrm{RF}_1$ for the exposed and not recovering class of females, but $\mathrm{RF}(t)$ increases as a function of time for the exposed recovering class:

$$\mathrm{RF}(t) = \min\left[1, \mathrm{RF}_1 + \frac{t-1}{T_R}(1 - \mathrm{RF}_1)\right]. \qquad (4)$$

An alternative way to view the population model is in the form of a matrix model (Caswell, 2001):

$$\mathbf{N}_{t+1} = \mathbf{P}\mathbf{N}_t, \qquad (5)$$

where $\mathbf{N}_t$ is a vector containing the number of animals in each age, sex, and exposure class at time t, and $\mathbf{P}$ is the population projection matrix that projects the number of animals at time t to the number at time $t + 1$. Elements of the matrix contain the parameters for age-, sex-, and exposure-class specific survival and fecundity. An advantage of this formulation is that analytic methods exist (see Caswell, 2001) to determine the stable age structure (i.e., expected proportion of animals in each class if the projection matrix parameters are constant through time), expected lifespan, and average age, all of which we used in the population simulation.

To initialize the model, we sampled each input parameter from a distribution (Table 1) estimated using methods detailed below. Given the sampled input parameters, the number of dolphins in each age and sex class was obtained by assuming the population was at its stable age distribution (i.e., by multiplying N_0 by the stable age proportions obtained from the dominant right eigenvector of the model population projection matrix given parameters for that year [Caswell, 2001]). The population was not assumed to be at equilibrium in year 0 and so, given the DD fecundity in the model, the population parameters were not constant over time; hence, the population would not be exactly at the stable age proportions. We undertook a small study (not shown) to investigate the difference between this transient population structure and the stable age structure and found it was negligible; therefore, we used the latter because it was available analytically and far faster to generate.

To evaluate injury, we simulated forward for 50 years under two scenarios: an oil spill (injured) scenario and a scenario in which there was no oil spill (baseline scenario, all animals are unexposed). We calculated three metrics of injury as described by Schwacke et al. (2017): maximum proportional decrease (MPD), difference between the two population trajectories when the injured trajectory is at its lowest point divided by the baseline; years to recovery (YTR), number of years required before the injured population trajectory reaches 95% of the baseline population trajectory; and lost cetacean years (LCY), sum of annual differences between baseline and injured population sizes.

We repeated this procedure for each of 10,000 iterations, each time simulating a set of realizations from the distributions of all inputs, projecting forward under injured and baseline scenarios and then calculating the injury metrics. The model was implemented in R version 4.0.5 (R Development Core Team, 2021).

Full model code and input data, as well as R Markdown documents describing how each component of the model was assembled and code to generate the output figures, are available at https://github.com/TiagoAMarques/CARMMHApapersSI and have been uploaded to GRIIDC (https://doi.org/10.7266/n7-rmv4-qx46).

Initial population size

No survey data existed from BB just before the DWH spill that could be used to estimate prespill population size in 2010; hence, other data sources had to be used to infer this quantity. Schwacke et al. (2017) used the average population abundance estimate from a Bayesian spatial capture-recapture (SCR) analysis of photo-ID data collected in BB from 2010 (postspill) to 2014, as reported by McDonald et al. (2017). The SCR analysis for the BB population has been updated by Glennie et al. (2021), who included additional data from later surveys, updated (likelihood based) analysis methods, and report estimates that are model averaged among different plausible models. Similar to McDonald et al.'s (2017) results, they found that population size in the study area increased from 2010 to 2012, likely as animals returned from peripheral areas outside of the photo-ID survey area but still within the bay, once disturbance related to cleanup operations subsided. We, therefore, took as the best estimate of prespill population size the mean of the estimated population size in mark-recapture primary periods 6 and 7 (midpoints 14 February 2012 and 15 April 2012). This is almost certainly an underestimate of prespill population size because the total population likely decreased in size between the time of the oil spill (April 2010) and primary periods 6 and 7. Oiling in BB was extensive (Nixon et al., 2016) and occurred outside the photo-ID survey area. We, therefore, assumed that all dolphins in BB, even if those that may have been outside of the photo-ID survey area during the first few primary periods, were exposed to oil.

We used a parametric bootstrap procedure to generate initial population size (N_0) values from the Glennie et al. (2021) SCR models. Under the assumption that the maximum likelihood estimates of the SCR models each followed a multivariate normal distribution, 10,000 parameter sets were generated from the fitted models. The number of samples from each model was in proportion to its model weight (Glennie et al., 2021). Each parameter set was used to derive estimates of population size in primary periods 6 and 7, and these were averaged yielding 10,000 replicate population sizes for N_0.

Age- and sex-specific baseline survival

Schwacke et al. (2017) obtained age- and sex-specific estimates of survival under "baseline" conditions with age at death data for stranded dolphins from five populations. In brief, Schwacke et al. (2017) fit a Siler (1979) competing risk function, which is based on the assumption that probability of surviving to a given age is the product of three competing risks: an exponentially

TABLE 1 Input parameters for the model of bottlenose dolphins in Barataria Bay

Parameter	Description	Point estimate (SD)	Distribution	Distribution parameters or summary	Data sources
N_0, $N_{nominal}$	Initial population size; also nominal population size	3045 (375.3)	Sampled from spatial capture recapture (SCR) outputs	95% CI 2720–3611	Glennie et al. (2021)
Age- and sex-specific baseline survival					
$S_{f,i}^a$, $S_{m,i}^a$	Age-specific annual survival for females and males of age i under a baseline scenario		Sampled from Siler model outputs		Schwacke et al. (2017)
Density-dependent baseline fecundity					
ASM	Age of (female) sexual maturity	8.6 (0.27)	Gamma	shape = 980.4, scale = 0.00875	Lacy et al., 2021; Mead & Potter, 1990
F_{max}	Maximum fecundity rate	0.34 (0.0133)	Beta-PERT	min = 0.33, mode = 0.34, max = 0.41, shape = 4.0	Schwacke et al. (2017)
$F_{nominal}$	Nominal fecundity rate	0.24 (0.0333)	Beta-PERT	min = 0.13, mode = 0.24, max = 0.33, shape = 4.0	Schwacke et al. (2017)
ρ	Shape parameter for density-dependent fecundity function	9.93 (2.95)	Shifted gamma	shape = 6.34, scale = 1.172, min = 2.5	Expert elicitation (EE)
Factors for reduction in survival					
S_1	Survival rate for first year postspill	0.82 (0.055)	Sampled from SCR outputs	95% CI 0.70–0.91	Glennie et al. (2021)
$S_{baseline}$	Baseline survival	0.94 (0.005)	Sampled from Siler model outputs, and P(marked|age)	95% CI 0.93–0.95	Schwacke et al. (2017)
P(marked|age)[a]	Proportion of animals marked as a function of age		Sampled from logistic model		Data supplied by L. Schwacke
Factors for reduction in fecundity					
R_1	Reproductive success rate for first year postspill	0.19 (0.004)	Beta	a = 1608, b = 6664	Kellar et al. (2017)
$R_{baseline}$	Baseline reproductive success rate	0.65 (0.007)	Beta	a = 3275, b = 1786	Kellar et al. (2017)
Factors for recovery					
PR	Probability of recovery to a baseline health state	0.183 (0.121)	Scaled beta	min = 0, max = 0.55, a = 1.17, b = 2.36	EE

[a] No details are provided in the table because parameter is a vector and hence not easily represented in a tabular format. Full distributions for these are available on the Github repository (https://github.com/TiagoAMarques/CARMMHApapersSI). P(marked|age) is the probability of being marked as a function of age.

decreasing risk due to juvenile factors, a constant risk experienced by all age classes, and an exponentially increasing risk due to senescent factors. Schwacke et al. (2017) used Markov chain Monte Carlo methods and 4000 posterior distribution combinations of the Siler model parameters to compute probability of a dolphin surviving to age x for $x = 1 \cdots 60$, independently for males and females. Following Schwacke et al. (2017), we used the 4000 resulting life tables as input values to the population dynamics model, randomly drawing with replacement a paired male and female lifetable for each population simulation.

Age at sexual maturity and reproductive senescence

The population model was based on the assumption that females younger than a certain age at time t cannot produce offspring in year $t + 1$; we refer to this as age at sexual maturity (ASM). Although Schwacke et al. (2017) used a static value (8 years), we made ASM an input parameter and estimated a distribution for ASM values through a meta-analysis. We searched the literature for studies that estimated ASM for bottlenose dolphins either by undertaking necropsies on bycaught or stranded animals and recording observations of ovarian scars or by recording first observation of a calf for known-age individuals. For the live dolphin studies, we assumed calves were observed immediately after birth, gestation time was 1 year, and females first became pregnant immediately following sexual maturity. Therefore, the estimate for ASM would be 1 year prior to the first observation of a calf. There are three species of bottlenose dolphins: *T. truncatus*, *aduncus*, and *australis*. Given that species may differ in ASM, we retained only articles relating to *T. truncatus*. We also excluded studies from captive animals because they may reproduce earlier (e.g., if food conditions are better) or later (if there is less social stimulus) than wild animals. Given these criteria, we retained only two studies for meta-analysis: a study based on necropsies of stranded animals from the U.S. central Atlantic Coast (Mead & Potter, 1990) and a study based on long-term monitoring (1982–2019) of 53 known-age females from an SB, Florida (Lacy et al., 2021). For the Mead and Potter (1990) data, a GAM was fitted with sexual maturity as a binary response, age was modeled as a thin-plate regression spline, and a logit link function was used. R. Wells supplied raw data (age of each female when she was first observed with a calf) from the analysis of Lacy et al. (2021), from which we calculated empirical mean and SE. To combine the results, we took an inverse variance weighted mean and matched resultant mean and variance with a corresponding gamma distribution for ASM. In each iteration of the simulation, ASM was obtained by sampling a random deviate from the gamma distribution and rounding to the nearest integer.

Schwacke et al. (2017) assumed that females at age 48 and older do not breed, based on data from SB. However, given the small number of animals of that age studied, we did not assume reproductive senescence in our model. Because there are very few animals of this age in the simulated populations, whether they are assumed to breed or not would have a negligible effect on population trajectories.

Density-dependent baseline fecundity

We used the modified form of the Beverton–Holt function as described by Schwacke et al. (2017), which does not require a specific assumption about carrying capacity, to model fecundity as a function of population size at time t:

$$f(N_t) = \frac{F_{max}}{1 + (\beta \times N_t)^\rho} \qquad (6)$$

where

$$\beta = \frac{1}{N_{nominal}} \times \left(\frac{F_{max} - F_{nominal}}{F_{nominal}} \right)^{\frac{1}{\rho}} \qquad (7)$$

and F_{max} is the maximum achievable fecundity rate, $F_{nominal}$ and $N_{nominal}$ are estimates of the fecundity and population size at a particular point in time, and ρ determines the shape of the function. Following the approach of Schwacke et al. (2017), we adopted beta-PERT distributions for F_{max} and $F_{nominal}$ and set $N_{nominal}$ as the N_0.

Schwacke et al. (2017) identified the DD shape parameter as being a major component of uncertainty in the injury metrics. Therefore, to improve specification of this parameter, we undertook a formal EE for ρ (described below).

Reduced survival factor

The SF for exposed classes was calculated as the population average survival measured in the first year after the spill, S_1, divided by population average baseline survival, $S_{baseline}$:

$$SF_1 = \frac{S_1}{S_{baseline}}. \qquad (8)$$

In practice, survival estimates for the entire population were not available after the oil spill. Estimates from the SCR analysis of photo-ID data were available, but because such studies can only be carried out on animals with distinctive markings on their dorsal fins, younger individuals are underrepresented in the photo-ID sample. Dolphins are not typically born with distinctive fin markings, but acquire them during their lifetime. Therefore, for calculating SF, we defined the population as being the population of marked animals, and we assumed that SF was the same across all age classes.

The estimate of S_1 was obtained from SCR analysis of BB dolphins by Glennie et al. (2021). We used estimated survival between the first and second primary sampling periods, 26 June 2010–12 November 2010, annualized, to represent first year postspill survival. This estimate was based on information from the entire photo-ID time series because survival is assumed to vary smoothly over time in the Glennie et al. (2021) models. Multiple models were fit, and degree of smoothness in survival over time varied among models. For the population model, we used 10,000 survival estimates, sampling from each model in

proportion to its model weight and so accounting for uncertainty in the smoothness.

Estimates of age- and sex-specific baseline survival were available from the Siler model. We had to average these in an appropriate way to produce an estimate of baseline survival for the cohort of the population that was marked. This required accounting for proportion of animals in each age and sex class and probability of an animal being marked as a function of age and sex. For the former, we used stable age and sex distribution obtained from the baseline population dynamics model for $t = 0$. For the latter, we undertook an analysis of data on presence and absence of marks for known-age dolphins. We analyzed age data from 54 dolphins in BB ranging in from 2.5 to 42 years with recorded presence or absence of distinctive marks, such as would be used in photo-ID surveys. We then used binary regression to estimate probability of being marked as a function of age. Preliminary modeling indicated no evidence for sex differences, so both sexes were analyzed together. An absence of very young animals made prediction difficult for these ages. Therefore, the models were constrained, so probability of being marked at age 0 was fixed at 0. A GAM was used with presence or absence of marking as the response variable, age as the explanatory variable modeled as a cubic regression spline, and a logit link function. From the fitted model, 10,000 parametric bootstrap resamples were generated (Appendix S2). The $S_{baseline}$ was calculated as the weighted average of age- and sex-specific baseline survival values. Weights were given by the expected proportion of dolphins in each age or sex class at t_0 multiplied by the estimated probability of being marked for the given class.

Reduced reproduction factor

Fecundity of reproductive age females in the year after the oil spill was assumed to be equal to the DD fecundity, $f(N_t)$, multiplied by a reduced reproductive success factor:

$$\text{RF}_1 = \frac{R_1}{R_{baseline}}, \tag{9}$$

where R_1 is reproductive success in the first year postspill and $R_{baseline}$ is a baseline reproductive success for dolphins not exposed to DWH oil. Following Schwacke et al. (2017), we derived reproductive success rates from results reported by Kellar et al. (2017) for dolphins from reference sites not exposed to oil ($R_{baseline}$) and from observed reproductive success in the first 3 years postspill at BB and Mississippi Sound (R_1). For both quantities, we had observations on number of trials, N, and number of successes, n. These naturally induced an observed proportion of successes: $p = n/N$. We used these to parametrize a beta distribution for reproduction success probability before and after the oil spill. The corresponding beta distribution has mean and SD that would correspond to a binomial distribution with N and p as observed. Reproductive success differed from fecundity in that the former was the probability a pregnant female successfully gives birth, whereas the latter was the probability a reproductive age female in year t gives birth to a calf in year $t + 1$.

Expert elicitation

To address uncertainty in two highly sensitive parameters for which direct empirical data did not exist, we conducted formal EEs. Expert elicitation is a structured process by which expert knowledge regarding an uncertain quantity is translated into a probability distribution. It has been applied in a range of scientific fields when there is urgent need for management decisions but data are lacking (Booth & Thomas, 2021; Martin et al., 2012). The parameters elicited were shape of DD fecundity function (ρ) and proportion of the population that would eventually recover to baseline survival (P_R).

We conducted two workshops, one with six experts in cetacean population biology and dynamics to address DD fecundity and the other with six experts in marine mammal health and physiology to address the recovery of dolphin health state. The design of the elicitation followed the methods of Booth and Thomas (2021) and is summarized below. Experts were informed of the purpose of the elicitation and invited to participate. Informed consent was obtained from all participants involved in the study (compliant with the framework outlined in title 45 U.S. Code of Federal Regulations 46.116).

The panels were designed such that a comprehensive coverage of opinions could be achieved (following guidance in Gosling, 2018), meaning expert knowledge spanned bottlenose dolphin ecology and the fields of population biology, demography, epidemiology, animal physiology, and veterinary science. Each group of experts was provided with an evidence dossier summarizing available background knowledge on DD relationships in mammalian populations (for ρ) and the analyses of temporal trends in prognoses scores for dolphins born prior to and following the DWH oil spill (for P_R). In discussion with the experts, the scope of the elicitation, definitions, and questions were clarified to ensure linguistic uncertainty was removed. Once questions were finalized, experts provided judgments through variable-interval methods for the quantity of interest. Experts used a shiny app to anonymously submit their judgments, which were subsequently fit as probability distributions with SHELF software (Gosling, 2018). All fitted distributions were anonymized and presented to the group, and each expert was invited to provide their rationale for their judgments. These rationales were discussed as a group to reach consensus of what would be a rational view of an impartial observer who had been party to all information presented and discussed (details in Astfalck et al., 2018).

The BB population is among 31 bay, sound, and estuary populations of common bottlenose dolphins in the northern Gulf of Mexico (NOAA, 2019). Experts agreed to the wording of questions and provided judgments to estimate the shape of DD response: "Given that the value of Rho will affect the shape of the DD response of fecundity in a bay, sound, and estuary bottlenose dolphin population, what do you judge to be the most appropriate value for Rho?" To estimate proportional recovery in survival probability of those affected by DWH oiling, the question was "Think about a Barataria Bay bottlenose

dolphin whose health has been impacted by the DWH oil spill and has a guarded, poor or grave prognosis in the first year or 2 after the DWH spill. Going forward, in the animal's lifetime, what is the probability that it has a good or fair prognosis?"

Because the second question was framed in terms of proportion of animals that may recover during their lifetime, we needed to determine average remaining lifespan of a dolphin alive at the time of the DWH spill to implement recovery in the model. The average remaining lifespan was used as T_R; that is, we assumed dolphins in the recovering class that are still alive T_R years after the oil spill will have made a full recovery and will have baseline levels of survival and fecundity. Average remaining lifespan was calculated as average total lifespan minus average current age—both of these quantities were in turn calculated from the population projection matrix at $t = 0$.

Population model sensitivity analyses

We conducted an uncertainty analysis to evaluate how uncertainty on each input parameter affected the injury metrics. For each parameter, we sampled from the input distribution and ran the model, holding all other parameters at their nominal value (i.e., point estimates in Table 1). We also conducted an elasticity analysis to evaluate proportional change in each injury metric resulting from a 1% increase in each variable (i.e., nominal value $\pm$ 0.5%) while holding all other variables at nominal values. The uncertainty analysis was useful because it pointed to which variables caused the greatest uncertainty in injury metrics. This uncertainty had two causes: variability in the input variable and sensitivity of the injury metric to variation in the input variable. The elasticity analysis was useful because it allowed us to isolate the second of these.

RESULTS

Temporal trends in health data

In the decade following the DWH spill, we conducted 262 assessments that included sufficient data for assignment of health prognosis score on dolphins in BB ($n = 171$) and SB ($n = 91$) (Appendix S1).

For dolphins alive when the spill occurred, the BB cohort had a higher prevalence of guarded, poor, or grave prognoses than the SB cohort ($p<0.0001$), and prevalence changed over time ($p = 0.02$) (Figure 2 & Appendix S1). The prevalence decreased in the first years following the spill (2013–2014), but then increased, reaching a maximum of 78% in 2018 (Figure 2a). Several disease conditions contributed to the guarded, poor, and grave scores in BB pre-2010 dolphins; lung disease weighed heavily in prognosis assignment (Smith et al., 2021). Other primary concerns included inflammation and anemia. In contrast, the SB pre-2010 cohort had few guarded and no poor or grave scores (Figure 2b), and although the prevalence of guarded prognoses was higher in 2016–2018
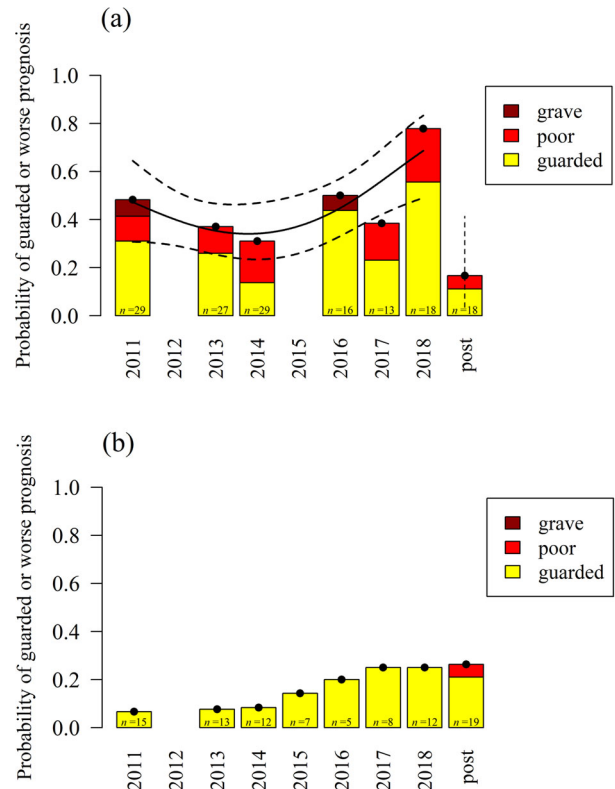

FIGURE 2 Probability of guarded, poor, or grave prognoses across years for cohorts of bottlenose dolphins born prior to the *Deepwater Horizon* oil spill in 2010 and prevalence for the cohort of dolphins born after 2010 (post, all years combined) in (a) Barataria Bay and (b) Sarasota Bay (fitted curves, predicted values; dashed-line curves, 95% CIs from generalized additive model with separate smoothing splines for sampling year [Barataria Bay $p = 0.02$, Sarasota Bay $p = 0.09$]; vertical dashed lines, postspill cohorts' 95% binomial confidence intervals)

compared with earlier years, there was no strong evidence of a temporal trend ($p = 0.09$). For post-2010 cohorts, there was no significant difference in prevalence of guarded, poor, or grave scores between BB and SB ($p = 0.19$).

Age of sexual maturity

From the necropsy data, the mean (SE) estimated age at which probability of being sexually mature was 50% was 7.76 years (1.51) (obtained via parametric bootstrap).

For the observational data from SB, mean (SE) age of first observing a calf was 9.60 years (0.27). This produced a mean ASM of 8.60 (0.27).

Combining the results from the two studies, we estimated an inverse variance weighted mean (SE) ASM of 8.57 years (0.27). Given that the estimated SE was so much larger in the estimate from the necropsy data than the SB live dolphin data, the combined result was very close to that from the SB data alone. The fit to a gamma distribution yielded a shape parameter of 980.4 and scale parameter of 0.00875 (Table 1).

Density-dependent fecundity

The population biology experts agreed that reproductive investment would be reflected in interbirth intervals and that as populations approach carrying capacity (K), resource limitations and increased risk of disease would start to affect populations. They considered it unlikely that DD effects would occur until a population gets relatively close to K, giving a concave DD fecundity function (i.e., with $\rho > 3$). Experts considered pregnancies in bottlenose dolphins were, energetically, relatively low cost and resource limitations would not be experienced until getting closer to what the habitat can support (>75% of K). Experts also noted that social structure in bottlenose dolphin populations might result in fecundity remaining high even as they approach K. Specifically, that social relationships and spending time in groups are beneficial to reproductive success (Mann et al., 2000; Wells, 2014). Therefore, at higher densities, reproductive females may be more resilient to resource limitations. The EE exercise produced a distribution for ρ with a mean of 9.94, which is higher than the estimate of 6 used by Schwacke et al. (2017) (Figure 3a). Higher values for ρ would allow a population to recover more quickly to K (Figure 3b).

Reduced survival and reproduction factors

The estimated SF for the first year following the spill was 0.868 (95% CI 0.737–0.964). The estimated RF for the first year was 0.694 (95% CI 0.438–0.875).

Proportion that recover

Experts for the health-recovery elicitation discussed the underlying conditions (particularly endocrine and pulmonary effects) that influence prognoses scores, processes that could inhibit recovery, and considered the likelihood that the types of disease conditions observed were fully recoverable or not. Specific cases of individual dolphins that had been sampled across multiple years, some of which showed improving condition and some of which did not (Smith et al., 2017), were discussed. The experts agreed the probability that the average dolphin would recover was low but that animals with guarded prognoses might be more likely to recover versus those with poor or grave prognosis. This resulted in a skewed distribution with some weight allowing for greater likelihood that animals might recover (Figure 4).

Population model outputs, uncertainty, and elasticity

Estimated MPD was 0.45 (95% CI 0.14–0.74) and would occur about 10 years postspill and take 35 years (95% CI 18–67) for 95% recovery (Figure 5). The LCY due to the spill was 30,993 (95% CI 6607–94,148). Median predicted population size in year 9 following the spill was 1816 (95% CI 893–2801). In comparison, SCR analyses provided an abundance estimate of 2099 (95% CI 1852–2505) based on photo-ID surveys conducted in
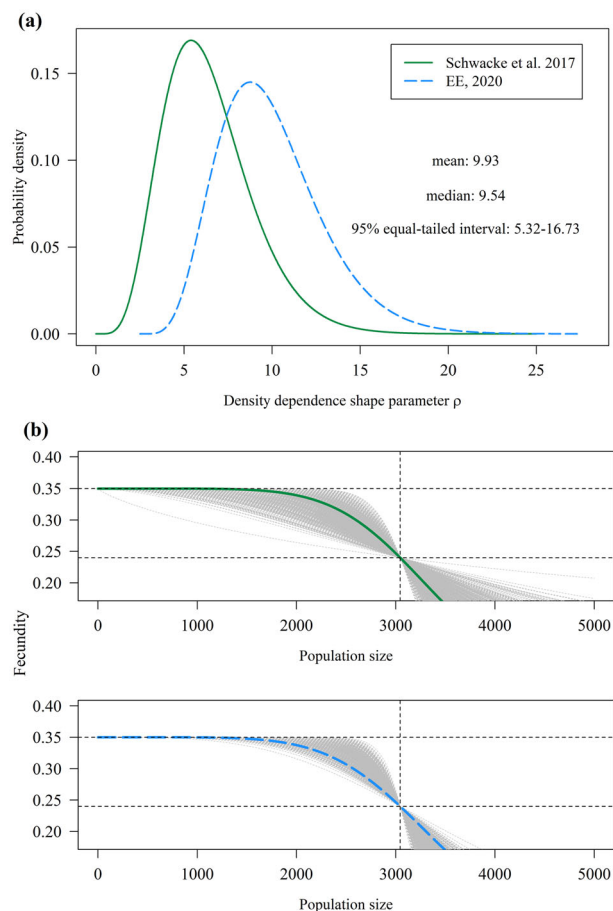
FIGURE 3 Probability distributions of the (a) consensus of the expert elicitation for density-dependence shape parameter, ρ, for a bay, sound, and estuary population of bottlenose dolphins compared with ρ values from Schwacke et al. (2017) and (b) corresponding mean function of fecundity versus population size for the Barataria Bay dolphin population obtained by sampling values of ρ from the distributions in (a), for which $N_{nominal} = 3045$ (dotted vertical line), $F_{nominal} = 0.24$, and $F_{max} = 0.34$ (dotted horizontal lines)
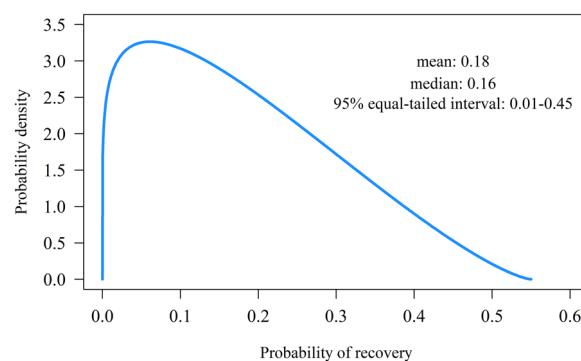
FIGURE 4 Elicited distribution for the probability of recovery over a lifetime for Barataria Bay dolphins exposed to *Deepwater Horizon* oil

2019 (Glennie et al., 2021). We used the analyses of Glennie et al. (2021) to derive N_0, but we did not use their SCR abundance estimates in subsequent years to fit population trajectory. In this sense, our population model prediction and the SCR estimate for 2019 can be considered independent.
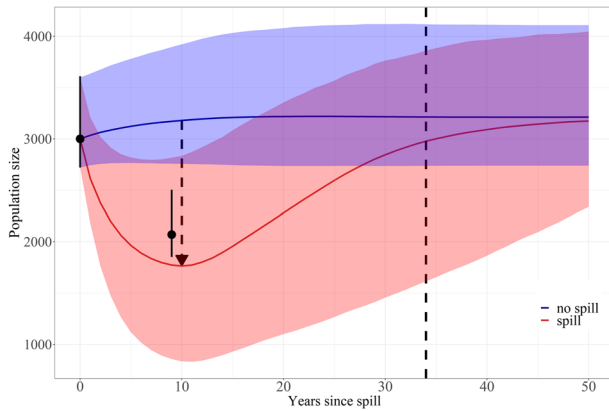
FIGURE 5 Median predicted size for the Barataria Bay dolphin population under baseline scenario (blue line) and *Deepwater Horizon* injury scenario (red line) from 10,000 simulations (shading, 95% CIs; dashed vertical line, years to recovery; dashed arrow, maximum proportional decrease; black dots, population abundance estimates from spatial capture-recapture (Glennie et al., 2021); solid vertical lines, 95% CIs). The year 0 estimate was used to initialize the population simulations, but the year 9 estimate was not used directly in the population model

Postspill survival (S_1) generated the greatest uncertainty in output metrics, particularly for LCY, and had relatively high elasticity (Appendix S3). This parameter estimate comes from the SCR analysis of photo-ID data (Glennie et al., 2021), and confidence intervals were quite wide (95% CI 0.70–0.91). This uncertainty in S_1 combined with its high elasticity meant that it was the biggest contributor to overall uncertainty in injury.

Baseline age- and sex-specific survival (S_f and S_m) had the highest elasticity values for LCY and YTR, although their contribution to uncertainty was lower than S_1 because these parameter estimates were more precise (e.g., SD on $S_{Baseline}$), which was essentially a weighted average of S_f and S_m values, was an order of magnitude lower than that of S_1. The one exception in elasticity was MPD, for which elasticity on S_1 was larger than S_f and S_m. In this case, S_1 was a more important determinant of MPD, which occurs close to the beginning of the postspill timeline, whereas baseline survival became more important in predicting injury metrics associated with time to recovery, YTR, and to a lesser extent LCY.

Remaining parameters had lower elasticities; consequently, contributions to output uncertainty came primarily from higher uncertainty in input parameters. Higher values of ρ elicited at the EE meant that this parameter had much lower elasticity compared with the Schwacke et al. (2017) model. Because a given level of variation in ρ produces smaller differences in DD population trajectories at high ρ than low ρ (Figure 3b), this parameter also contributed far less to overall uncertainty than in the Schwacke et al. (2017) model.

DISCUSSION

Like most large mammals, dolphins display traits of slow-living species (Stearns, 1992) in that they mature late, produce a single offspring at a time, invest multiple years in care of each offspring, and are long-lived. Therefore, acute mortality alone, particularly affecting adult females, can produce population

declines that require years for recovery. Our results suggest that lower survival is a continuing problem for BB dolphins alive during the DWH spill and has led to a continued population decline even 9 years postspill. Even when these animals are replaced by those born after the spill, it will take decades to recover baseline population numbers. Empirical data support our model's predictions. The most recent abundance estimate (Glennie et al., 2021) is 30.9% lower (95% CI 22.1–39.5) than prespill (Figure 5). Similar declines were seen in killer whales (*Orcinus orca*) in years following the *Exxon Valdez* oil spill; neither of the affected pods recovered within 16 years (Matkin et al., 2008). Because studies following the *Exxon Valdez* spill were unable to conduct hands-on health assessment or even recover carcasses for necropsy, links among oil exposure, physiological effects, survival, and population decline were circumstantial. In the absence of physiological data, small population size and demographic factors, rather than chronic toxic effects, were proposed as underlying mechanisms for a lack of recovery (Esler et al., 2018). The DWH spill, while tragic, offered an opportunity to directly study toxic effects and better understand long-term consequences of oil and associated chemical exposure on cetacean health, survival, and reproduction—factors that ultimately drive population recovery. Our findings of continued chronic and potentially progressive disease in dolphins exposed to oil provide insight into why it can take so long for animals to recover.

We applied a series of innovative analysis approaches, including a formal EE, to data collected in the decade following the DWH spill to refine input parameters for a multiclass structured population model. The time series data confirmed that some disease conditions became chronic and progressed over time, and expert input provided interpretation for the consequences. Specifically, lung disease, presumably from inhalation and aspiration of oil, did not improve, and in some cases, worsened over time (Smith et al., 2021). This finding was key to experts' conclusion that probability of an exposed individual ever fully recovering to baseline health was low. The experts also agreed that reproductive recovery would likely follow the trajectory for survival, consistent with findings that reproductive success is closely related to maternal health (Barratclough et al., 2021). Lack of recovery of individuals exposed to oil implies population recovery must rely on emergence and sexual maturation of a healthier cohort almost entirely founded from postspill progeny. Fortunately, our data support that health of postspill progeny is comparable to health of similar age dolphins from a comparison site not exposed to DWH oil. However, an investigation of functional immune responses, which used samples from the same dolphin health assessments, suggests exaggerated Th2 responses in unexposed BB dolphin offspring, like those in the prespill cohort, which may increase the susceptibility to infectious disease (De Guise et al., 2021). It is possible that insufficient time has passed for clinical symptoms to manifest in dolphins born postspill.

Differences in injury metrics we obtained from our model versus those of Schwacke et al. (2017) were not large and primarily arose as a consequence of additional knowledge gained in the intervening period (Appendix S4). The new estimate for shape of the DD fecundity function from the EE was larger

than estimates used by Schwacke et al. (2017), which had been based on a model previously fitted for DD pup survival in gray seals (*Halichoerus grypus*). Larger values of ρ will produce smaller injury metrics because they will allow populations to recover more quickly to carrying capacity. However, this was balanced by a slightly higher estimate for ASM (8.6 vs. 8.0 years used by Schwacke et al., 2017) and a lower estimate for S_1, both of which serve to slightly increase injury metrics. Our updated input information and refined analyses validated and reinforced the original assessment that the DWH spill caused tens of thousands of lost cetacean years, that approximately halved the dolphin population at its low point, and that it will take 3–4 decades for the population to recover without effective restoration efforts.

The ultimate goal of an NRDA is to determine the appropriate type and amount of restoration needed to return negatively affected natural resources, including animal populations, to their original condition. Our model construct is well-suited to integrate potential changes in vital rates over time and could be easily adapted to evaluate proposed restoration plans and what influence they might have on the recovery trajectory. The current focus for DWH restoration in Louisiana is a project involving sediment diversion from the Mississippi River to rebuild wetlands (Cornwall, 2021). The planned project will substantially reduce salinity and has been forecast to decrease dolphin annual survival by 34% after implementation (Garrison et al., 2020). This level of impact is clearly unsustainable and would result in population depletion rather than restoration.

Much can be learned about oil spill effects in marine mammals from the DWH injury assessment and subsequent research. Foremost, we have shown that previously understudied inhalation and aspiration exposure pathways are important and that pulmonary effects are likely to be long-lasting, decrease survival and reproduction, and ultimately impede population recovery. Our results demonstrate the importance of effective response, long-term investment to monitor recovery, effective restoration options, and monitoring of restoration performance for adaptive management.

From this, we provide recommendations for managing future chemical spills that affect cetacean habitat. For oil spill response efforts, we recommend the identification of methods to eliminate surface oil quickly to minimize inhalation and aspiration of oil. The application of dispersants is controversial because of perceived toxicity risk and increased risk for respiratory effects (Alexander et al., 2018). Further research is needed to evaluate benefits and risks of dispersants, specifically for mammals.

More generally for NRDA injury assessment, likely long-term health implications and life history traits must be considered in assessing injury and allocating restoration resources. The population model we described not only provides metrics of long-term population-level effects, but also demonstrates a framework for inclusion of expert knowledge for parameters with high uncertainty due to lack of empirical data. Although our model was developed for dolphins, it is applicable for other species with similar life history traits.

Our sensitivity analyses showed that uncertainty in post-spill survival was the largest determinant of uncertainty in population injury. For long-lived species, small changes in adult survival create large changes in population trend (i.e., high elasticity [Caswell, 2001]). Obtaining precise estimates of time-varying survival is difficult at the start of capture-recapture studies when the catalog of known individuals is small. Therefore, in future similar situations when the study is only initiated after the spill, initial field efforts to build a robust photo-ID catalog are critical and great effort should be dedicated to follow as many individuals as possible from the start.

Finally, increased susceptibility of individuals, which we found comes from injuries that are likely to be lifelong conditions, must be monitored long term, so that restoration projects can be adapted over time, and should be considered in restoration planning. In the near-term, restoration efforts for cetaceans must focus on reduction of other stressors (e.g., other pollutants, marine debris, and fishing gear) to prevent worsening health and population risk because there are no proven medical interventions to address persistent disease conditions associated with oil exposure. However, the DWH spill has highlighted the need to develop such intervention strategies, and this should be a priority research area.

ORCID

Lori H. Schwacke https://orcid.org/0000-0003-0434-5332
Tiago A. Marques https://orcid.org/0000-0002-2581-1972
Len Thomas https://orcid.org/0000-0002-7436-067X
Cormac G. Booth https://orcid.org/0000-0002-5804-3455
Brian C. Balmer https://orcid.org/0000-0001-9734-4042
Patricia E. Rosel https://orcid.org/0000-0002-6481-3989
Ryan Takeshita https://orcid.org/0000-0003-2748-8803
Randall S. Wells https://orcid.org/0000-0001-9793-4181

REFERENCES

Alexander, M., Engel, L. S., Olaiya, N., Wang, L., Barrett, J., Weems, L., Schwartz, E. G., & Rusiecki, J. A. (2018). The *Deepwater Horizon* oil spill coast guard cohort study: A cross-sectional study of acute respiratory health symptoms. *Environmental Research*, *162*, 196–202.

Astfalck, L., Cripps, E., Gosling, J., Hodkiewicz, M., & Milne, I. (2018). Expert elicitation of directional metocean parameters. *Ocean Engineering*, *161*, 268–276.

Barratclough, A., Gomez, F. M., Morey, J. S., Meegan, J. M., Parry, C., Schwacke, L., Jensen, E. D., & Smith, C. R. (2021). Biochemical and hematological biomarkers of reproductive failure in bottlenose dolphins *Tursiops truncatus*. *Diseases of Aquatic Organisms*, *144*, 197–208.

Booth, C., & Thomas, L. (2021). An expert elicitation of the effects of low salinity water exposure on bottlenose dolphins. *Oceans*, *2*, 179–192.

Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation*. Sunderland, MA: Sinauer Associates.

Cornwall, W. (2021). Unleashing big muddy. *Science*, *372*, 334–337.

De Guise, S., Levin, M., Jasperse, L., Herrman, J., Wells, R. S., Rowles, T., & Schwacke, L. (2021). Long-term immunological alterations in bottlenose dolphin a decade after the *Deepwater Horizon* oil spill in the Northern Gulf of Mexico: Potential for multigenerational effects. *Environmental Toxicology and Chemistry*, *40*, 1308–1321.

Esler, D., Ballachey, B. E., Matkin, C., Cushing, D., Kaler, R., Bodkin, J., Monson, D., Esslinger, G., & Kloecker, K. (2018). Timelines and mechanisms of wildlife population recovery following the *Exxon Valdez* oil spill. *Deep Sea Research Part II: Topical Studies in Oceanography*, *147*, 36–42.

Garrison, L. P., Litz, J. & Sinclair, C (2020). Predicting the effects of low salinity associated with the MBSD project on resident common bottlenose dolphins (Tursiops truncatus) in Barataria Bay, LA. NOAA technical memorandum NMFS-SEFSC-748.

Glennie, R., Thomas, L., Speakman, T., Garrison, L., Takeshita, R. & Schwacke, L. (2021). Estimating spatially-varying density and time-varying demographics with open population spatial capture-recapture: a photo-ID case study on bottlenose dolphins in Barataria Bay, Louisiana, USA. arxiv.org/abs/2106.09579.

Gosling, J. (2018). *SHELF: The Sheffield Elicitation Framework*. Springer.

Kellar, N. M., Speakman, T. R., Smith, C. R., Lane, S. M., Balmer, B. C., Trego, M. L., Catelani, K. N., Robbins, M. N., Allen, C. D., Wells, R. S., Zolman, E. S., Rowles, T. K., & Schwacke, L. H. (2017). Low reproductive success rates of common bottlenose dolphins *Tursiops truncatus* in the northern Gulf of Mexico following the Deepwater Horizon disaster (2010–2015). *Endangered Species Research*, *33*, 143–158.

Lacy, R., Wells, R., Scott, M., Allen, J., Barleycorn, A., Urian, K., & Hofmann, S. (2021). Assessing the viability of the Sarasota Bay community of bottlenose dolphins. *Frontiers in Marine Science*, *8*, 1851.

Litz, J. A., Baran, M. A., Bowen-Stevens, S. R., Carmichael, R. H., Colegrove, K. M., Garrison, L. P., Fire, S. E., Fougeres, E. M., Hardy, R., Holmes, S., Jones, W., Mase-Guthrie, B. E., Odell, D. K., Rosel, P. E., Saliki, J. T., Shannon, D. K., Shippee, S. F., Smith, S. M., Stratton, E. M., Tumlin, M. C., Whitehead, H. R., Worthy, G. A. J. & Rowles, T. K. (2014). Review of historical unusual mortality events (UMEs) in the Gulf of Mexico (1990–2009): Providing context for the multi-year northern Gulf of Mexico cetacean UME declared in 2010. *Diseases of Aquatic Organisms*, *112*:161-175.

Mann, J., Connor, R. C., Barre, L. M., & Heithaus, M. R. (2000). Female reproductive success in bottlenose dolphins (*Tursiops* sp.): Life history, habitat, provisioning, and group-size effects. *Behavioral Ecology*, *11*, 210–219.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Choy, S., McBride, M., & Mengersen, K. (2012). Eliciting expert knowledge in conservation science. *Conservation Biology*, *26*, 29–38.

Matkin, C. O., Saulitis, E. L., Ellis, G. M., Olesiuk, P., & Rice, S. D. (2008). Ongoing population-level impacts on killer whales *Orcinus orca* following the *Exxon Valdez* oil spill in Prince William Sound, Alaska. *Marine Ecology Progress Series*, *356*, 269–281.

McDonald, T. L., Hornsby, F. E., Speakman, T. R., Zolman, E. S., Mullin, K. D., Sinclair, C., Rosel, P. E., Thomas, L., & Schwacke, L. H. (2017). Survival, density, and abundance of common bottlenose dolphins in Barataria Bay (USA) following the *Deepwater Horizon* oil spill. *Endangered Species Research*, *33*, 193–209.

Mead, J., & Potter, C. (1990). Natural history of bottlenose dolphins along the central Atlantic coast of the United States. In S. Leatherwood, & R. Reeves (Eds.), *The bottlenose dolphin*. San Diego, CA: Academic Press.165–195.

Nixon, Z., Zengel, S., Baker, M., Steinhoff, M., Fricano, G., Rouhani, S., & Michel, J. (2016). Shoreline oiling from the Deepwater Horizon oil spill. *Marine Pollution Bulletin*, *107*, 170–178.

NOAA. (2019). Common bottlenose dolphin (*Tursiops truncatus truncatus*): Northern Gulf of Mexico Bay, Sound, and Estuary Stocks. National Oceanic and Atmospheric Administration Marine Mammal Stock Assessment Report. https://media.fisheries.noaa.gov/dam-migration/170_tm258_bottlenosebse.pdf.

R Development Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schwacke, L. H., Smith, C. R., Townsend, F. I., Wells, R. S., Hart, L. B., Balmer, B. C., Collier, T. K., De Guise, S., Fry, M. M., Guillette, L. J., Jr., Lamb, S. V., Lane, S. M., McFee, W. E., Place, N. J., Tumlin, M. C., Ylitalo, G. M., Zolman, E. S., & Rowles, T. K. (2014). Health of common bottlenose dolphins (*Tursiops truncatus*) in Barataria Bay, Louisiana, following the deepwater horizon oil spill. *Environmental Science & Technology*, *48*, 93–103.

Schwacke, L. H., Thomas, L., Wells, R. S., McFee, W. E., Hohn, A. A., Mullin, K. D., Zolman, E. S., Quigley, B. M., Rowles, T. K., & Schwacke, J. H. (2017). Quantifying injury to common bottlenose dolphins from the *Deepwater Horizon* oil spill using an age-, sex- and class-structured population model. *Endangered Species Research*, *33*, 265–279.

Siler, W. (1979). A competing risk model for animal mortality. *Ecology*, *60*, 750–757.

Smith, C. R., Rowles, T. K., Hart, L. B., Townsend, F. I., Wells, R. S., Zolman, E. S., Balmer, B. C., Quigley, B., Ivancˇic′, M., McKercher, W., Tumlin, M. C., Mullin, K. D., Adams, J. D., Wu, Q., McFee, W., Collier, T. K., & Schwacke, L. H. (2017). Slow recovery of Barataria Bay dolphin health following the *Deepwater Horizon* oil spill (2013–2014), with evidence of persistent lung disease and impaired stress response. *Endangered Species Research*, *33*, 127–142.

Smith, C. R., Solano, M., Lutmerding, B. A., Johnson, S. P., Meegan, J. M., Le-Bert, C. R., Emory-Gomez, F., Cassle, S., Carlin, K., & Jensen, E. D. (2012). Pulmonary ultrasound findings in a bottlenose dolphin *Tursiops truncatus* population. *Diseases of Aquatic Organisms*, *101*, 243–255.

Smith, C., Rowles, T., Gomez, F., Colegrove, K., Takeshita, R., Zolman, E., Balmer, B., Wells, R., Townsend, F., & Schwacke, L. (2021). Evidence of poor pulmonary health in bottlenose dolphins following the *Deepwater Horizon* oil spill. International Oil Spill Conference Proceedings 2021. Allen Press.

Stearns, S. C. (1992). *The evolution of life histories*. Oxford; New York: Oxford University Press.

Venn-Watson, S., Colegrove, K. M., Litz, J., Kinsel, M., Terio, K., Saliki, J., Fire, S., Carmichael, R., Chevis, C., Hatchett, W., Pitchford, J., Tumlin, M., Field, C., Smith, S., Ewing, R., Fauquier, D., Lovewell, G., Whitehead, H., Rotstein, D., …, & Rowles, T. (2015). Adrenal gland and lung lesions in Gulf of Mexico common bottlenose dolphins (*Tursiops truncatus*) found dead following the deepwater horizon oil spill. *PLoS One*, *10*, e0126538.

Wells, R. (2014). Social structure and life history of bottlenose dolphins near Sarasota Bay, Florida: Insights from four decades and five generations. Yamagiwa, J. & Karczmarski, L., *Primates and cetaceans: field research and conservation of complex mammalian societies, Primatology Monographs*. (Pp. 149–172). Tokyo: Springer.

Westerholm, D. A. & Rauch S. D. (2016). Deepwater Horizon oil spill: Final programmatic damage assessment and restoration plan and final programmatic environmental impact statement. Retrieved from http://www.gulfspillrestoration.noaa.gov/restoration-planning/gulf-plan.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.