Health of North Atlantic right whales *Eubalaena glacialis* over three decades: from individual health to demographic and population health trends

Rosalind M. Rolland¹,*, Robert S. Schick², Heather M. Pettis¹, Amy R. Knowlton¹, Philip K. Hamilton¹, James S. Clark³, Scott D. Kraus¹

¹John H. Prescott Marine Laboratory, New England Aquarium, Boston, MA 02110, USA
²Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Fife KY16 9LZ, UK
³Nicholas School of the Environment, Duke University, Durham, NC 27708, USA

ABSTRACT: Marine mammals are faced with increasing challenges from environmental fluctuation, climate change, and disturbances from human activities. Anthropogenic mortalities have been well documented, but it is difficult to assess the sub-lethal effects of disturbance on the fitness of marine wildlife, and to distinguish these impacts from natural variations in health and reproduction. Here, we used photographic data on body and skin condition, blowhole cyamids, and rake marks, to evaluate the health of North Atlantic right whales *Eubalaena glacialis* from 1980 to 2008. We applied a hierarchical Bayesian model to these data to estimate the underlying continuous health status of individuals, demographic groups, and the population to characterize health patterns and temporal trends. Visual health scores (scaled from 0 to 100) from 48,560 sighting events were used to estimate the health of 622 identified right whales on a monthly basis. Health in most whales fluctuated between 70 and 90, and health scores of <60 were observed in whales in poor condition. Health varied by sex, age-class and reproductive state, with the greatest annual variability occurring in actively reproducing females. Calving females had significantly higher health scores than non-calving females, and a steep deterioration in population health coincided with a dramatic decline in calving from 1998 to 2000. Health in all demographic groups and the population declined over the 3 decades of observations. Given the inevitable data gaps that occur in most marine wildlife research, modeling advances such as the one presented here offer a promising approach to assess the complex interactions between biology, ecology, and sub-lethal anthropogenic disturbance on marine mammals.

KEY WORDS: North Atlantic right whale · Health · Fitness · Body condition · Reproduction · Bayesian model

INTRODUCTION

Maintenance of robust wildlife populations and recovery of endangered marine mammals depends upon the fitness of individuals within those populations. Naturally occurring environmental fluctuation, human-induced ecological changes, and direct impacts of human activity can have profound effects on the health and survival of individuals (Doney et al. 2012, Poloczanska et al. 2013, van der Hoop et al. 2013). Where these pressures are widespread and severe, negative effects on the health of individuals may lead to decreased growth, reproductive success, recruitment and survival, potentially resulting in biologically significant consequences at the population level (Caswell et al. 1999, Fujiwara & Caswell 2001,
Monitoring the health of wildlife is uniquely challenging in marine ecosystems, where sightings of individuals are sporadic and data are difficult to obtain. This is especially true for most large whales because they are only partially visible during brief surface intervals, and logistically difficult to sample (Rolland et al. 2007a, Hunt et al. 2013). Nevertheless, several approaches have been developed to assess relative body condition of large, free-swimming cetaceans, including aerial photogrammetry, blubber thickness measurements using ultrasound, and visual assessment from photographs (Perryman & Lynn 2002, Pettis et al. 2004, Rolland et al. 2007b, Fearnsbach et al. 2011, Miller et al. 2011, 2012, Bradford et al. 2012). Studies of harvested large whales have confirmed the links between blubber thickness (or mass) as a measure of energetic reserves, prey availability, and reproduction in large whale species (Lockyer 1986, Miller et al. 2011, 2012, Williams et al. 2013). Additionally, the prevalence and appearance of skin lesions has been related to both infectious disease organisms and environmental factors (Van Bressem & Van Waerebeek 1996, Wilson et al. 2000, Kiszka et al. 2009, Hart et al. 2012). Furthermore, non-lethal anthropogenic impacts including fishing gear entanglements and propeller marks from vessel collisions have been evaluated using characteristic scars or the presence of gear (Knowlton & Kraus 2001, Bradford et al. 2009, Knowlton et al. 2012). Thus, creative approaches have yielded data on the relative health and fitness of some species over the past decade. However, it has been very difficult to determine whether impacts on the health of individuals translate into significant biological consequences at the population level.

The western NA right whale is one of the most thoroughly studied populations of large whales, owing to 35 yr of continuous monitoring of these individually identifiable whales (Kraus & Rolland 2007). This endangered population currently numbers just over 500 individuals (Pettis & Hamilton 2014). Population growth has been hindered by mortalities from fishing gear entanglements and collisions with vessels, along with depressed reproductive rates (Kraus et al. 2005, Browning et al. 2010). Extensive data on these whales is amassed in the North Atlantic Right Whale Catalog (http://rwcatolog.neaq.org). Researchers believe that a combination of natural and anthropogenic stressors have negatively affected the health and vital rates of right whales, but the relative impact of different factors has been difficult to quantify (Kraus & Rolland 2007).
Schick et al. (2013b) used the extensive data available on individual NA right whales to develop a hierarchical Bayesian (HB) model to infer the links between health status, movement and survival of individual right whales. This state-space model incorporated 3 decades of sightings and associated data, survey effort, life history data (including calving), visual health scores, and entanglement events. This model was further modified by Schick et al. (2016) to address model assumptions in the first iteration that affected some of the health estimates made by the model. Here, we expand the scope of this previous work by using individual right whale health estimates to investigate variations in the health of different demographic groups, and at the population level, and to characterize changes in health over 3 decades. The specific objectives of this study were to: (1) use individual whale health profiles to create health profiles for different demographic groups or life history stages, (2) explore the temporal patterns and magnitude of variation of estimated health in different demographic groups, (3) examine temporal trends and variation in health at the population level, and (4) explore the relationship between estimated health of reproductive females and fecundity. The health observations on which this model was based incorporated influences of normal biology, environmental factors and human disturbance impacting health (e.g. fishing gear entanglement, vessel collisions). The goals of this study were to develop a quantitative understanding of health across the NA right whale population over 3 decades, explore the link between health and reproduction, and provide the foundation for future work on the effects of environmental variation and sub-lethal anthropogenic disturbance on fitness, fecundity and survival.

MATERIALS AND METHODS

General approach

In this study, we analyzed posterior estimates of health from the HB model detailed in Schick et al. (2013b). In particular, the model provided monthly estimates of latent health for individual right whales, with a data model to link this latent health state to photographic observations of health (Pettis et al. 2004). Broadly speaking, the analyses conducted here proceeded as follows: (1) we fit the model from Schick et al. (2013b) to visual health assessment (VHA) data to generate posterior estimates of health on a monthly time scale for each individual right whale, (2) we tallied health for specific demographic groups and a representative population group, (3) post-hoc statistical analyses were conducted on these data to test for differences in decadal periods and between different demographic groups, and (4) the influence of health on reproduction was assessed by comparing the health of calving vs. non-calving females, and by exploring the health of demographic groups and the population during intervals of low and high yearly calving rates. The following sections describe in detail the data used and analyses conducted.

Right whale databases

This model incorporated sightings and photo-identification data extending back to 1935, but the majority of data were from 1980 (when dedicated right whale surveys commenced in the Bay of Fundy, Canada) through 2008 (when the health data analysis was mostly completed). Photographs of callosity patterns on the rostrum along with unusual scars, markings or pigmentation on the body, flippers and flukes were used to identify individuals (Kraus et al. 1986). Data were obtained from the North Atlantic Right Whale Identification and Sightings Databases (Right Whale Consortium 2011). These databases provided effort and sightings data (date, time, latitude, longitude) along with corresponding individual whale-based data on age (or estimated age based on first sighting), sex, calving history, movements, mortality, and photographic evidence of health and anthropogenic impacts (i.e. fishing gear entanglement) for identified right whales.

Visual health assessment data (VHA)

Where possible, a full body series of lateral photographs from the tip of the rostrum to the flukes were evaluated to provide data for health assessment. Otherwise, all available images of adequate quality and photographic angle were examined and scored for as many parameters as possible. Four health parameters were assessed: body condition, skin condition, rake marks forward of the blowholes (combined scores for right and left sides), and presence and density of orange cyamids along the blowhole margins (Fig. 1) (Pettis et al. 2004). Parameters were scored on an ordinal scale of 1–3 (1 = poor, 2 = fair, 3 = good) or 1–2 (1 = poor, 2 = good); note that this scale is inverted from Pettis et al. (2004). All individual whale...
images were grouped by year and sighting region (as described in the North Atlantic Right Whale Sightings Database), and this collection of images was called a ‘batch’. Batches were assigned a single score for every health parameter that could be evaluated. For batches in which (a) parameter(s) changed (e.g. skin condition was good during early sightings within the batch but changed to poor by later sightings or vice versa), the final score was used. All photographs and images were evaluated by a single experienced right whale biologist, thus minimizing any potential inter-observer bias. Pettis et al. (2004) conducted a comparison of inter-researcher scoring for the VHA method using Kappa’s weighted test for

Fig. 1. (a) A North Atlantic right whale (EG# 3911) in good health observed on 10 February 2010 (Photo credit: Florida Fish and Wildlife Conservation Commission, NOAA Permit No. 775-1875). (b) The same right whale, observed on 15 January 2011, in poor health after a severe fishing gear entanglement that resulted in her death (Photo credit: Georgia Department of Natural Resources, NOAA Permit No. 932-1905/MA-009526). Poor body condition was evident from concavity in the dorsal profile in the post-blowhole area (denoted by white arrow), skin lesions and widespread orange cyamid coverage (yellow circles), orange cyamids along the margins of the blowholes (white circle), and rake marks anterior to the blowholes (yellow arrow). A white fishing line can be seen exiting the margin of the lips next to the yellow circle on the left.
agreement, and found strong agreement amongst experienced researchers for all the assessed parameters except rake marks, for which there was moderate agreement.

**Individual whale health profiles**

Following methods described in Schick et al. (2013b, 2016), we fitted an HB model to the visual health data for each individual right whale in the population, and generated monthly estimates of health. In terms of the length of each individual's estimated health time series, for whales with a known time of death (when the carcass was recovered and identified), we included health estimates from the month of first sighting to the month of death. However, most right whales simply disappeared from the sighting record and the time of death was not known. In these cases, we examined posterior estimates of health and survival, and no whale that had a health value below 32.5 remained alive. Therefore, we excluded any further monthly health estimates after a whale's health score declined below 32.5.

**Health trends in demographic groups**

Individual whale health profiles were aggregated, on an annual basis, into 7 demographic groups by sex, age-class (juvenile vs. adult) and reproductive state (for females). Demographic groups were defined as (1) young juveniles (1−2 yr), in which the health effects of the transition from weaning to independent feeding at 12−13 mo (Hamilton et al. 1995) are captured, (2) older juveniles (3−8 yr), and (3) adults (≥9 yr, or the year prior to the first calving event for females <9 yr). Health estimates for individual whales were averaged monthly to create summary health profiles for each of these demographic groups. To examine the effects of the calving cycle on changes in health, reproductive females were further subdivided based on calendar year: pregnant = 1 yr before a known calving event, lactating = the year of calving and nursing, resting (recovery year) = 1 yr after lactation, and ‘available to be pregnant’ = all years after resting until the next gestational year (hereafter referred to as ‘available’) (see Supplement 1 at www.int-res.com/articles/suppl/m542p265_suppl.pdf for more information on right whale calving cycles). The ‘available’ category excludes females with 3 yr calving intervals since they would be pregnant, and thus not available, in the year after their resting year. Therefore, the ‘available’ group contains females with lengthened inter-calving intervals (≥4 yr), and a female can remain in this category for multiple years (Fig. S1 in Supplement 1). Only intervals between known calving events were incorporated into the model. Nulliparous adult females were excluded only from analyses of health in reproducitively active females (since they were not reproducitively competent), but we compared their health estimates to the other adult female groups to see if their health was different from females that were successfully calving.

Health trends in different demographic groups were compared in the context of known life history characteristics of right whales (e.g. body condition loss in females as lactation progresses) using health estimates summarized monthly, annually, and over the entire 28 yr study period. Additionally, mean health of demographic groups was compared by decade (1980−1989, 1990−1999, 2000−2008).

**Population-level health trends**

Health estimates from adult males and older juveniles were averaged on a monthly basis to represent health trends at the population level. Adult males were included to reflect ambient environmental conditions without the energetic and physiological demands of calving experienced by adult females. Older juveniles were also incorporated as they experience the elevated energetic requirements of growth, and their survival is essential for recruitment into the population. Population health trends were characterized using health estimates summarized monthly, annually, and over the entire 28 yr study period. Population health was also compared by decade, as for the demographic groups.

**Health and reproductive success**

As capital breeders, it has been well established that successful reproduction in large baleen whales requires accumulation of adequate energetic reserves to support pregnancy and especially lactation (Lockyer 1986, Miller et al. 2011, 2012, Williams et al. 2013). Our previous work showed that fluctuations in body condition with the calving cycle were detectable using visual assessment (Pettis et al. 2004, Rolland et al. 2007b). Here, we explored the relationship between estimated health and reproductive success by comparing the annual mean health scores of rest-
ing and available females that transitioned to pregnant in the following year and successfully calved, versus corresponding females from those years that did not subsequently calve. These analyses included data from 1988–2008 (and relevant data from calving events pre-1988 and post-2008), as it required several years of sightings after 1980 to determine if a female was reproductively mature, resting or available. Additionally, a period of at least 3 yr was required after the end of the time series to determine if a female calved, or not, based on inter-calving intervals in right whales (Kraus et al. 2007) (Fig. S1 in Supplement 1). Since 1984, extensive aerial and shipboard surveys in the southeastern US calving ground have captured the majority of calving events in this population (Browning et al. 2010). Reproductive success was determined by observing the female in close association with a calf in multiple sightings (Knowlton et al. 1994).

**Health in low and high calving years**

Steep declines in calving rates occurred in this right whale population during two 3 yr intervals in the 1990s (24 and 11 calves were born in the periods 1993–1995 and 1998–2000, respectively). Calf counts during these times were significantly lower than expected with normal stochastic variation (Kraus et al. 2007). The causes of reduced calving in these years remains unknown; nevertheless, these events presented the opportunity to retrospectively determine if there was a coinciding decline in the modeled health of different demographic groups and the population. Estimated health of the demographic groups and the representative population group were compared during the intervals of depressed calving rates (1993–1995 and 1998–2000) and the 3 yr period (between 2001 and 2003) when the highest number of calves were born (n = 71). In addition, we included a comparison of health between 2004–2006, since there was an unexplained decrease in the modeled health in most groups, although only a slight decline in calving numbers (21.3 calves yr⁻¹) compared to the mean for that decade (24 calves yr⁻¹).

**Modeling approach**

Details of the modeling framework can be found in Schick et al. (2013b). Briefly, in this previous work, we built an HB model and fit it to NA right whale sightings data to make inferences on individual whale health, movement between habitats, and survival. The sightings component of the model is outlined in Schick et al. (2013b). Whilst we did not have detection probabilities from each track-line in the sense of the full DISTANCE sampling approach, we did build a complete sightings model that provided posterior estimates of sightability for each individual whale; thus, we explicitly accounted for sighting heterogeneity.

In the current study, we focused on the results from the health component of the HB model. The process model for health is updated at a monthly time step; current health (i.e. at time $t$) is a function of health at the previous time step ($t-1$), age, and process error. The ordinal VHA parameters were linked to latent health using multinomial logit functions; these functions relied on expert opinion-based priors (see Schick et al. 2013b, their Appendix 1 for details). Latent health was estimated on a continuous scale of 0–100, and the health estimate initializes with the first sighting of the individual whale. Starting values in the Markov Chain Monte Carlo (MCMC) for health were based on the average condition for each of the 4 VHA parameters.

We modified the HB model from Schick et al. (2013b) in 3 ways: (1) setting a more restrictive prior on the process error variance, (2) setting health proposals within the Gibbs sampler based on data from H. Pettis (unpubl. data), and (3) altering how we imputed missing data for body and skin condition. The restrictive prior was used because whales appeared to recover more quickly from more serious injury, and this slows the rate of health change. For the health proposals, i.e. how far from the current value a new value can be proposed, we settled on a value of 5. Finally, we altered the method for missing data imputation from Schick et al. (2013b) by implementing a linear interpolation scheme between observations, and by only imputing missing data within ±6 mo of a sighting of an individual whale. This 6 month window was informed by the mean time required to visually detect changes in body condition in North Atlantic right whales (H. Pettis unpubl. data). We further refined this to limit backwards imputation of health in lactating females to avoid the situation where a VHA value from early in a lactating year would be imputed backwards into the pregnancy year. Together these changes have the effect of slightly reducing month-to-month variability in health estimates. Details of the first and third model refinements above are given in Schick et al. (2016).
Details of the computation and model-fitting process are given in Schick et al. (2013b). Briefly, we fit the model to data using a Gibbs sampler using MCMC techniques (Clark 2007). We ran the model for 50000 iterations within the Gibbs loop, discarded these values as burn-in, and ran the model for another 50000 iterations. Standard MCMC techniques were used to determine if convergence was reached using the superdiag (Tsai & Gill 2012) package in R (R Core Team 2014).

Statistical analysis

Statistical analyses were performed using R (R Core Team 2014). Standard descriptive statistics, including means, SD, and medians, were calculated. Very large sample sizes and normal data distributions permitted use of parametric tests. The data used in these tests are the monthly estimates of health for each individual right whale; for example, in one year for one whale, there would be 12 separate measurements included in the analyses. Statistical tests included Welch’s 2-sample \( t \)-test for comparison of means, and 1-way ANOVA with Tukey post-hoc tests for comparisons of health between demographic groups, decades and time periods. We tested differences in adult female health as a function of calving success using Welch’s \( t \)-test and a linear mixed effects model (Bates et al. 2014, 2015). The fixed effects portion of the model was the health value in the available year as a function of pregnancy status (pregnant/not pregnant) in the following year. The random effects portion was year. The significance level for all statistical analyses was \( \alpha = 0.05 \).

RESULTS

Individual whale health

Data used in this study included VHA scores for 11931 batches of images encompassing 48560 sighting events. Each batch included between 1 and 66 images of each individual whale, and encompassed 1–3 mo of sightings. The VHA data included observations of body condition (\( n = 8963 \)), skin condition (\( n = 13397 \)), rake marks (\( n = 9315 \)) and blowhole cyamids (\( n = 7441 \)). There were 79 cases in which a visual health parameter changed score within the batch, and the final score in the batch was used for the entire batch. Time series of health profiles were generated for 622 individual right whales. Estimates of individual health varied widely, ranging from 0 to 93.4, with most healthy whales fluctuating between 70 and 90. In lactating and resting females, lower health scores largely reflected normal body condition loss while nursing a calf, and recovery of condition post-weaning (see example in Fig. 2). Scores lower than 60 were mostly seen in whales in poor condition, such as those with severe fishing gear entanglements as reported in Schick et al. (2013b). Credible intervals around health estimates were wider for whales with sparse sighting histories.

The time series of observed and estimated health for an adult female right whale named Staccato (EG# 1014) illustrates the model output for individual whales on which this study was based (Fig. 2). Staccato had a 25 yr sighting history, starting in 1974 and ending with her death in 1999 after a vessel collision. During a total of 133 sightings, Staccato was photographed in most of the right whale critical habitats. She had 6 calves and 3 minor fishing gear entanglements. Her health fluctuated from 70 to 85 until a steep decline occurred shortly before her death. Based on necropsy results showing a partly healed mandibular fracture, it is possible that this whale was first struck by a vessel 1 to 2 wk before being killed by a second fatal vessel strike (Moore et al. 2004).

Demographic group health

The time series of mean estimated health data for the 7 demographic groups (Fig. 3), along with yearly sample sizes for each cohort (see Table 1 for total sample sizes for each group) are displayed starting in 1988 to allow known-age whales (especially juveniles) to populate the database (there were almost no known-age whales in 1980). Health profiles differed by sex, age-class and, in females, reproductive status. Juveniles and, to a lesser extent, adult males maintained higher health scores throughout the study period (Figs. 3a & 4). Adult females had the largest and most consistent short-term variability in estimated health, mostly due to loss of body condition from the energetic demands of the calving cycle (Fig. 3b). Available and pregnant females had similar health patterns, while resting and lactating females maintained lower health scores throughout the study (Figs. 3b & 4). The lowest mean health scores (60.7) were observed in resting females in 1997 (Fig. 4); the year before calving rates plummeted. Mean health in nulliparous females (\( n = 9 \)) did not differ from the
Fig. 2. Time series (1974−1999) of the observed and estimated health for a female North Atlantic right whale (EG# 1014), which died in 1999 following a vessel collision. Visual health data used to estimate health (top 4 panels) were color-coded according to the rank score of each batch of sightings, and included body condition and rake marks (green = good, orange = fair, purple = poor), and skin condition and blowhole cyamids (green = good, orange = poor) (Pettis et al. 2004). In the ‘health’ panel, modeled health estimates (solid line) and the 95% Bayesian credible interval for the posterior that represents the uncertainty around estimated health (gray ribbon) are shown, and compared with the health of the population (dashed line). Estimated health scores range from 1 to 100, with lower scores indicating worse health. The ‘anomaly’ panel shows negative deviations of individual health from population health. Bottom panels show calving (‘calf’; light gray in gestation year, black in lactation year) and entanglement events (‘entgl’; green = minor injuries), which are factors potentially impacting health, but were not explicitly included as data in the model. The gray line before the symbol denotes the time-frame within which the entanglement event occurred (last seen without gear or scarring from the entanglement event). Fishing gear entanglement was documented when photographed whales were carrying gear, or upon detailed examination of all sightings of a whale for characteristic wrapping scars resulting from previous encounters with gear (Knowlton et al. 2012)
other 4 adult females groups (data not shown). All groups except for pregnant and lactating females had an obvious decline in mean health during the period 1998−2000 coinciding with the lowest 3 yr interval for calf production (gray bars in Fig. 3). A second health decline was seen in the mid-2000s in most groups. Declining health over the entire study period was seen in all demographic groups (Figs. 3 & 4).

Estimated health data by demographic group over the entire study period is summarized in Fig. 5 (since data from all years are combined, the smaller sample sizes prior to 1988

Table 1. Mean (±SD) estimated health scores for the representative population group and the demographic groups by decade. Adult females are categorized into pregnant, lactating, resting and available (to be pregnant), and juveniles are divided into young (1−2 yr) and old (3−8 yr) groups. Sample sizes (n) are for the entire study period (1980−2008). ns: periods that were not significantly different (Tukey’s HSD); p << 0.001 for all groups.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>622</td>
<td>79.8 ± 9.1</td>
<td>75.2 ± 12.2</td>
<td>72.5 ± 12.6</td>
<td>1589_{2.66301}</td>
</tr>
<tr>
<td>Adult males</td>
<td>184</td>
<td>80.2 ± 9.0</td>
<td>75.3 ± 12.3</td>
<td>73.2 ± 12.5</td>
<td>1554_{2.74553}</td>
</tr>
<tr>
<td>Young juveniles</td>
<td>292</td>
<td>80.9 ± 8.4</td>
<td>75.1 ± 12.3</td>
<td>72.4 ± 12.7</td>
<td>2868_{2.63941}</td>
</tr>
<tr>
<td>Old juveniles</td>
<td>255</td>
<td>80.9 ± 8.6</td>
<td>75.6 ± 12.2</td>
<td>72.2 ± 12.9</td>
<td>2989_{2.68777}</td>
</tr>
<tr>
<td>Pregnant</td>
<td>150</td>
<td>77.7 ± 7.5_{ns}</td>
<td>78.1 ± 7.1_{ns}</td>
<td>73.2 ± 10.5</td>
<td>1672_{2.4947}</td>
</tr>
<tr>
<td>Lactating</td>
<td>156</td>
<td>75.3 ± 9.9</td>
<td>73.4 ± 9.0</td>
<td>70.4 ± 9.7</td>
<td>103.8_{2.4490}</td>
</tr>
<tr>
<td>Resting</td>
<td>156</td>
<td>77.2 ± 7.3</td>
<td>72.1 ± 10.3</td>
<td>70.9 ± 10.2</td>
<td>163.9_{2.3929}</td>
</tr>
<tr>
<td>Available</td>
<td>158</td>
<td>78.9 ± 5.0</td>
<td>73.5 ± 11.6</td>
<td>71.2 ± 12.9</td>
<td>124.3_{2.5097}</td>
</tr>
</tbody>
</table>
have a smaller effect on this analysis). The relationship between health patterns in different groups was very similar to that described on an annual basis above, and the mean health scores from all years combined differed significantly between groups (Tukey’s HSD; $F_{6,192} = 15.6$, p << 0.001).

**Population-level health**

The representative population health summary (based on combined health estimates for adult males and older juveniles) by month (1988–2008) is shown in Fig. 6. A decreasing health trend was apparent starting in the early 1990s, along with a multi-year decline starting around 1996 that coincided with the steep decline in calving rates (Fig. 6). Population health scores rebounded in 2000, but never reached the level of the health scores observed at the beginning of the time series. A very slight dip in health was also observed in 2004–2005 (Fig. 6).

**Health over 3 decades**

Mean estimated health for the population and the demographic groups declined significantly over the 3 decades (1980–2008) of the study period (Table 1). Health estimates were highest in the 1980s, and the lowest in the 2000s. Mean health differed significantly by decade in the population and all demographic groups except pregnant females, in which health in the 1980s and 1990s did not differ significantly (Table 1).
Reproduction and health

Although there were differences by year, overall, the annual mean health scores of resting and available females that transitioned to pregnant (n = 240) were significantly higher than for females that did not become pregnant in the following year ($t_{7470} = 4.6507$, $p << 0.001$, n = 368) (Fig. 7). The vast majority of females that became pregnant and calved ('successful' females) had health estimates >70. There were only 2 years (2004 and 2005) in which the mean health of successful females was below 70 (min. = 67). Based on this, the health threshold for successful reproduction appears to be ≥67. Results from the linear mixed effects model (see Supplement 2 at www.int-res.com/articles/suppl/m542p265_supp.pdf for details) also found that successful females were healthier (Table S1), and indicated support for including year as a random effect term (Table S2). In 1997 and 1998, the intercept was below 0 and confidence intervals did not include 0, with intercept values of −2.95 and −3.33 for 1997 and 1998, respectively. This pattern was also true for 2005−2007; intercept values were −5.82, −4.54, and −4.5. Esti-
mates of the intercept for year were negative in these 5 years, which indicates that overall health values were lower (Fig. S2).

Health in low and high calving years

Mean health estimates of the population and the demographic groups were compared during the two 3 yr intervals of lowest calving (1993−1995 and 1998−2000), the highest 3 yr calving interval (2001−2003), and between 2004−2006, when estimated health decreased, but calving rates did not decline significantly (Table 2). Health at the population level and in adult males differed significantly in all 4 time periods (Table 2). The lowest health for the population, adult males, and both juvenile groups was in 1998−2000, corresponding with the worst years for calving (1998−2000). The highest health was observed during 1993−1995.

In contrast, all female groups had the lowest mean health scores from 2004 to 2006 (Table 2). Pregnant and available females showed similar patterns of health differing in all time periods, with the highest health scores in the period from 1993 to 1995. Mean health of lactating females was similar during all periods except 2004−2006. In resting females, mean health was significantly lower in both 1998−2000 and 2004−2006. Health estimates for lactating and resting females were likely influenced by small sample sizes during years of low calf production.

DISCUSSION

In this study, a HB model was used to estimate underlying continuous health states at the level of individual right whales, demographic groups, and the population. The long-term study on NA right whales provided an extensive data set on visual health encompassing 48,360 sighting events of 622 individual whales collected over 3 decades. The results show that health varied by sex, age-class, and reproductive state, with actively reproducing females showing the greatest range in health values and the largest annual variability. The lowest esti-
Mated health scores were in resting and lactating females, reflecting normal loss of body condition due to the high energetic demands of lactation in right whales (Miller et al. 2011, Miller et al. 2012, Fortune et al. 2013). In contrast, juveniles maintained the highest plane of health, followed by adult males. A health threshold for successful reproduction (≥67) was suggested based on the lowest mean health scores for females that successfully calved. Available and resting females that transitioned to pregnant and calved in the subsequent year had higher mean health scores than corresponding females that did not calve (with variability by year), further demonstrating a link between estimated health and reproduction. In addition, a population-wide deterioration in health from 1998 to 2000 coincided with a drastic decline in calving rates, suggesting that factors influencing health at the population level were responsible for suppressed reproduction during this period. Finally, model results showed that health in all demographic groups and at the population level declined over 3 decades.

Recovery of the NA right whale population has been extremely slow (growth rates have varied from −0.02 to 2.5% over 3 decades) in sharp contrast to southern right whale populations that are growing at 7–8% yr⁻¹, and currently comprising tens of thousands of individuals (Kraus et al. 2007, Browning et al. 2010). While anthropogenic mortalities have slowed growth (Caswell et al. 1999, Fujiwara & Caswell 2001, Kraus et al. 2005), reduced reproductive rates and extremely variable annual calving numbers are also important contributing factors (Kraus et al. 2007, Browning et al. 2010). It is unknown whether this impaired reproduction is due to intrinsic (biological) or extrinsic (environmental and/or anthropogenic) factors, or a combination of both. While there is evidence that genetic factors may be chronically depressing reproductive success (Frasier et al. 2007), this does not explain the acute periods of extremely low calving (e.g. 1998–2000). Climate-change drivers related to the right whale’s calanoid copepod prey may have influenced calving rates (Meyer-Gutbrod & Greene 2014). There is also evidence that acoustic disturbance from underwater...
vessel noise and the effects of non-lethal fishing gear entanglements are impacting levels of physiologic stress (Rolland et al. 2012) and fecundity in these whales (A. Knowlton unpubl. data). Therefore, it is likely that both intrinsic and extrinsic factors play a role in reduced calving rates. As the ocean regions inhabited by North Atlantic right whales are much more impacted by human activity and ‘urbanization’ compared to those inhabited by southern hemisphere right whales (Kraus & Rolland 2007), multiple sub-lethal disturbances are probably inhibiting the recovery of this species. However, distinguishing between the effects of natural environmental variation and anthropogenic disturbance is difficult, especially when multiple stressors are co-occurring with normal fluctuations in health-related parameters (i.e. body condition loss during lactation).

In light of these myriad stressors, results from this model provide a retrospective synopsis of health patterns in NA right whales that reflects effects of natural stressors (lactation), environmental variability, and anthropogenic factors that influenced the visual health data on which the model was based. Changes in underlying health in response to varying ecological conditions may be detectable before impacts on fecundity and survival are apparent. Thus, health measures may provide a more sensitive indicator before large-scale demographic changes are detectable (Moore 2008). For example, changes in body condition in North Atlantic fin whales *Balaenoptera physalus* have been linked to prey abundance and pregnancy rate (Williams et al. 2013). In this study we did not attempt to parse out the sub-lethal effects of existing anthropogenic influences (e.g. fishing gear entanglement and vessel strikes) on estimated health, which will be explored in future work (see below). Nevertheless, demographic groups that may be uniquely vulnerable to disturbances were identified; resting and lactating females are particularly susceptible to factors affecting the quantity and quality of prey available (including possible effects of habitat disturbance) due to their depleted blubber reserves. In addition, the health trajectories for the demographic groups and the population can be used as a reference to assess the impact of additional future anthropogenic disturbances, which may have a greater effect on whales already compromised by environmental and anthropogenic factors affecting population fitness and resilience. Thus, establishing a quantifiable link between vital rates (reproduction) and health enhances the ability to predict the population consequences of a variety of sub-lethal anthropogenic stressors.

This model also presents a new approach to explore hypotheses to explain health declines, and to link changing health to reproductive success. Changes in health in different demographic groups, and during periods of good or poor calving success, can point to different underlying environmental factors or etiologic agents. The population-wide decrease in health from the year 1998 to 2000 points to broad-scale factors affecting all whales, such as compromised body condition associated with insufficient quality or quantity of prey (Greene & Pershing 2004, Greene et al. 2008, Meyer-Gutbrod & Greene 2014). It is possible that only higher quality individuals (i.e. individuals with higher fitness) were able to reproduce during this interval. However, there were likely other issues affecting reproduction during the period 1993–1995, since a similar health decline was not detected. Results of the linear mixed effects model examining the relationship between health of females and calving success also found a significant effect of year. In marine mammals there are a variety of agents that can cause reproductive failure that might not be detected visually, including harmful algal blooms producing marine biotoxins such as domoic acid (Brodie et al. 2006), and infectious organisms (reviewed in Van Bressem et al. 2009). Moreover, model output also showed significant health deterioration in all demographic groups and the population from 2004 to 2006 without a corresponding decrease in calving rate. During this interval, the health decline was driven partly by compromised skin condition scores primarily in females and young juveniles. This finding would lead to consideration of causes of skin lesions that do not impact reproduction as the etiologic factor affecting health during these years. Interestingly, a dramatic increase in right whale mortality, primarily from vessel collisions, was seen during this same interval (Kraus et al. 2005), and whether, or not, this was related to compromised health is unknown, but worthy of future investigation. In summary, this study provided insights into how changes in estimated health can be influenced by different visual health assessment parameters at different times, and it also showed that not all visual health declines were associated with impacts on reproduction.

Output from this model may be biased or limited by several factors. Body condition was probably underrepresented in health estimates because it can rarely be assessed in aerial images, and the model did not differentiate by survey platform. Because aerial surveys predominate in habitats where right whales are found in the winter and spring, body condition data is
mostly from summer and fall surveys, thus a decline in body condition in the early part of the year would be less likely to be reflected in health estimates. Likewise, behaviors (such as skim-feeding or sub-surface feeding) seen in certain habitats (e.g. Cape Cod Bay) preclude body condition evaluation, because of changes in the dorsal body profile caused by an arched back and open mouth. As a result, some of the visual health data from these habitats will be under-represented in the model. Additionally, a segment of the population use currently unidentified habitats at certain times of the year (Hamilton et al. 2007), and, given the lack of health data from these areas, this model would not capture any environmental conditions or other factors impacting health related to the use of these areas. Similarly, reproducing females are sighted frequently in the year of calving and lactation, but much less often during gestation and resting years (Brown et al. 2001), thus, observational health data along with the specifics of variation in health are more limited for these 2 female cohorts. It is also unknown whether unhealthy whales may be more, or less, likely to be sighted, which could influence the study results. Additionally, small sample sizes for certain demographic groups in some years affected some of the analyses presented here. For example, there were very small numbers of pregnant, lactating and resting females and young juveniles during the years of depressed calving, which undoubtedly impacted some of the comparisons. There were also smaller sample sizes at the beginning of the dedicated right whale surveys in the early 1980s that likely influenced the results of analyses from this decade. Also, the process model for health very likely under-represented all of the internal and external factors that contributed to health, including potential covariates such as prey abundance and climatic fluctuation (e.g. North Atlantic Oscillation). Finally, the potential for cumulative impacts on health from multiple stressors was not addressed in this model iteration.

In future work, this model framework could be expanded to incorporate additional data sets related to environmental and anthropogenic disturbance and right whale health. Our next objective is to quantify the effects of non-lethal fishing gear entanglements and vessel strikes on health. Schick et al. (2013b) reported that the lowest health scores were observed for whales with severe fishing gear entanglements, pointing to the need to investigate the sub-lethal impacts of entanglements of varying severity on individual health and vital rates. As 83% of NA right whales have been entangled in fishing gear at least once, and 15.5% of the population is entangled annually (Knowlton et al. 2012), it is crucial to understand whether this perturbation results in population level impacts. Additionally, approximately 14% of right whales have either been killed by vessel collisions or bear scars from ship encounters (A. Knowlton unpubl. data), therefore analysis of the health effects of non-lethal vessel collisions is also needed. Furthermore, through exploring the linkages between health and survival of individuals, improved estimates of population size and growth rates should be feasible. The existing model could also be used to explore the link between habitat use patterns and health to determine whether whales are more or less vulnerable to added disturbance in certain areas. Incorporation of data on prey abundance may clarify the relationships between food resources, body condition, and reproduction, and could help distinguish the effects of prey availability from anthropogenic factors. A growing body of underwater acoustic monitoring data collected in right whale habitats could be incorporated to investigate the relative impacts of low frequency underwater noise (Clark et al. 2009, Hatch et al. 2012). These additional variables could be modeled at the level of individual whales, demographic groups, and the population, as in the present study. Finally, photographic observations of body condition and entanglements have been used to evaluate health in other cetacean species (Bradford et al. 2009, Bradford et al. 2012), thus it should be feasible to extend this model to other well-studied populations of marine mammals.

Marine ecosystems will continue to be altered by anthropogenic disturbances and climate change, confronting wildlife with an evolving scenario of challenges. This modeling approach can provide insights into risk factors for both individual whales and populations, thereby helping to determine effective management or mitigation options. Longitudinal, individual-based studies are critical to provide the foundational data for this approach, and close collaboration between modelers and knowledgeable biologists is essential for development of accurate models. Identifying the current health status of populations may assist in predicting the impact of added anthropogenic disturbance, and special concern could be afforded to populations already experiencing a declining health trend. Given the difficulty of maintaining long-term field programs, and the inevitable data gaps that occur in marine research, modeling advances such as the one outlined here offer a promising method to assess the complex interactions between biology, ecology, and anthropogenic impacts on marine mammals.

Marine ecosystems will continue to be altered by anthropogenic disturbances and climate change, confronting wildlife with an evolving scenario of challenges. This modeling approach can provide insights into risk factors for both individual whales and populations, thereby helping to determine effective management or mitigation options. Longitudinal, individual-based studies are critical to provide the foundational data for this approach, and close collaboration between modelers and knowledgeable biologists is essential for development of accurate models. Identifying the current health status of populations may assist in predicting the impact of added anthropogenic disturbance, and special concern could be afforded to populations already experiencing a declining health trend. Given the difficulty of maintaining long-term field programs, and the inevitable data gaps that occur in marine research, modeling advances such as the one outlined here offer a promising method to assess the complex interactions between biology, ecology, and anthropogenic impacts on marine mammals.
Acknowledgements

This work was primarily supported by the US Office of Naval Research (ONR) Awards N000141010614 (S.K.), N000141210286 (L. Thomas, J. Harwood, C. Harris, and R.S.), and N000141210274 (E. Fleishman). This work also received funding from the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland) and their support is gratefully acknowledged. MASTS is funded by the Scottish Funding Council (grant reference HR09011) and contributing institutions. We are grateful to the members of the Population Consequences of Disturbance working group for their intellectual contributions, helpful comments and insight, and we thank Dr. Michael Wiese and Dana Belden for their support for this research and related efforts. Special thanks to Dr. Robert Kenney for providing the sightings and survey effort data. We wish to express our deepest gratitude to the North Atlantic Right Whale Consortium for access to images and data in the North Atlantic Right Whale Identification and Sightings Databases. This study would not have been possible without generous data sharing by the members of the North Atlantic Right Whale Consortium, an exemplary example of scientific cooperation to forward the conservation of North Atlantic right whales. Finally, we thank 4 anonymous reviewers whose thoughtful comments greatly improved this manuscript.

LITERATURE CITED


Brotzke LD, Rehstock GA (2014) Climate change increases reproductive failure in Magellanic penguins. PLoS ONE 9:e85602


Hickie BE, Ross PS, Macdonald RW, Ford JK (2007) Killer whales (Orcinus orca) face protracted health risks associ-
ated with lifetime exposure to PCBs. Environ Sci Technol 41:6613–6619


Schick RS, Kraus SD, Rolland RM, Knowlton AR and others (2013b) Using hierarchical Bayes to understand movement, health, and survival in the endangered North Atlantic right whale. PLoS ONE 8:e64166

Schick RS, Kraus SD, Rolland RM, Knowlton AR and others


Editorial responsibility: Scott Shaffer, San Jose, California, USA

Submitted: May 19, 2015, Accepted: November 8, 2015
Proofs received from author(s): December 9, 2015