Oecologia (2017) 183:353–365
DOI 10.1007/s00442-016-3774-3

PHYSIOLOGICAL ECOLOGY - ORIGINAL RESEARCH

Effectiveness of baseline corticosterone as a monitoring tool for fitness: a meta-analysis in seabirds

Graham H. Sorenson[1] · **Cody J. Dey**[2] · **Christine L. Madliger**[1,2] · **Oliver P. Love**[1,2]

Received: 29 June 2016 / Accepted: 8 November 2016 / Published online: 22 November 2016

Abstract Many ecosystems have experienced anthropogenically induced changes in biodiversity, yet predicting these patterns has been difficult. Recently, individual behavioural and physiological measures have been proposed as more rapid links between environmental variation and fitness compared to demographics. Glucocorticoid hormones have received much attention given that they mediate energetic demands, metabolism, and foraging behaviour. However, it is currently unclear whether glucocorticoids can reliably predict environmental and fitness-related traits and whether they may be useful in specific groups of taxa. In particular, seabirds are a well-studied avian group often employed as biomonitoring tools for environmental change given their wide distribution and reliance on large oceanic patterns. Despite the increase in studies attempting to link variation in baseline corticosterone (the primary avian glucocorticoid) to variation in fitness-related traits in seabirds, there has been no comprehensive review of the relationship in this taxon. We present a phylogenetically controlled systematic review and meta-analysis of correlative and experimental studies examining baseline corticosterone as a predictor of fitness-related traits relevant to predicting seabird population health. Our results suggest that, while variation in baseline corticosterone may be a useful predictor of larger-scale environmental traits such as overall food availability and fitness-related traits such as reproductive success, this hormone may not be sensitive enough to detect variation in body condition, foraging effort, and breeding effort. Overall, our results support recent work suggesting that the use of baseline glucocorticoids as conservation biomarkers is complex and highly context dependent, and we suggest caution in their use and interpretation as simplified, direct biomarkers of fitness.

Keywords Glucocorticoids · Systematic review · Cort-fitness · Biomarker · Food availability · Seabirds

Introduction

Ocean ecosystems have experienced extensive changes in biodiversity and species abundance as a result of anthropogenic impacts (Hoegh-Guldberg and Bruno 2010). Unfortunately, isolating the precise mechanisms underlying these changes has not always been straightforward (Halpern et al. 2008). Increasingly, wide-ranging marine vertebrates are being employed as useful and biologically relevant bio-monitoring tools of both fine- and large-scale variation in oceanic conditions (Hussey et al. 2015). Seabirds (penguins—Sphenisciformes, petrels/albatrosses—Procellariiformes, gannets and cormorants—Pelecaniformes, and gulls and auks—Charadriiformes; Fig. 1) represent an increasingly well-studied and diverse sub-group of predators, and are often considered prime biomonitoring tools given their wide distribution and reliance on oceanographic patterns across large spatial areas (Piatt and Sydeman 2007). Specifically, because oceanic environments

---

Communicated by Indrikis Krams.

Electronic supplementary material The online version of this article (doi:10.1007/s00442-016-3774-3) contains supplementary material, which is available to authorized users.

✉ Graham H. Sorenson
graham.sorenson@gmail.com

[1] Department of Biological Sciences, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada

[2] Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada

are changing rapidly, yet are so large and complex to study, seabirds have the potential to act as very practical and sensitive biomarkers to changes in oceanic patterns, food supply and location, and temperature regime shifts for a number of reasons (Frederiksen et al. 2007; Parsons et al. 2008). First, seabirds range over very large areas of temperate, tropical and polar oceans during different stages of their life-history (e.g. breeding, migration and wintering; Weimerskirch et al. 2014). Second, because many seabird species are central place foragers, a result of colonial breeding (Orians and Pearson 1979), researchers also often have the capacity to monitor individuals and populations to

Fig. 1 Diversity of seabird species studied worldwide. **a** Black-legged kittiwake (*Rissa tridactyla*), S. Descamps **b** african penguin (*Spheniscus demersus*), K. Hick **c** thick-billed murre (*Uria lomvia*), G. Sorenson **d** northern fulmar (*Fulmarus glacialis*), M. Mallory **e** antarctic petrel (*Thalassoica Antarctica*), S. Descamps **f** tufted puffin (*Fratercula cirrhata*), K. Elliott

determine causal links between large-scale oceanographic change and individual fitness and/or population health (Satterthwaite et al. 2012; Descamps et al. 2013). Third, seabirds are experiencing significant population declines with 28% of species listed as globally threatened and 10% near threatened (BirdLife International, Croxall et al. 2012). Finally, many of these declines are thought to be both directly and indirectly related to anthropogenically induced changes (e.g. Bodey et al. 2014). As such, it is important to ensure biomonitoring tools reflect the impacts of environmental variation on factors important to population demographics.

Traditionally, examining how environmental variation impacts the health of colonially breeding species has been accomplished via the long-term monitoring of population demography (Gaston et al. 2012; Descamps et al. 2013). Indeed, demographic studies have been important for inferring broad-scale linkages between oceanic changes and seabird colony trends (Paleczny et al. 2015). However, given that trends often show long temporal lags between environmental variation and population responses (Ewers and Didham 2006), they can be less effective for predictive, proactive or rapid-response management endeavours (Satterthwaite et al. 2012). In response to this information gap, integrative ecology has seen a surge in the use of field analytical techniques aimed at connecting finer-scale, individual-level mechanisms (e.g. behaviour, stable isotopes, immune system metrics, energetic physiology) as biomarkers of larger-scale environmental change with the hope of ultimately predicting population responses before they occur (Ozgul et al. 2010; Moody et al. 2012; Cooke et al. 2013; Ramos et al. 2014; Hussey et al. 2015). By focusing on specific traits that can reveal how environmental variation is impacting individuals, it is hypothesized that researchers can extrapolate these mechanisms to the population level and proactively determine how and why a given population may be changing (Madliger and Love 2016). For example, Harding et al. (2007) studied behaviour in relation to food availability in Common Murres (*Uria algae*) and demonstrated that time budgets could indicate large-scale food availability within certain spatial ranges. The authors suggest that by establishing the link between prey species and murre foraging, seabird behaviour can ultimately be used as a tool for estimating aspects of ocean productivity. In addition, physiological traits have recently been employed to examine a diversity of responses to the environment across taxa (Cooke and O’Connor 2010; Chown and Gaston 2015), and conservation physiologists have increasingly focused their efforts on traits associated with energetic management (Madliger and Love 2015).

Glucocorticoids (often referred to as ‘stress hormones’ for their role in the acute stress response; Benowitz-Fredericks et al. 2008), such as corticosterone (CORT) in birds, play an important regulatory role at baseline levels. Importantly, they provide a central function in allostasis by mediating variation in energetic demand, glucose management and foraging/feeding behaviour (Astheimer et al. 1992; Ricklefs and WIlkelski 2002; Landys et al. 2006; Angelier et al. 2008). Given that baseline glucocorticoids (GCs) should be sensitive to changes in the environment that impact allostatic management (McEwen and Wingfield 2010; Madliger and Love 2014, 2015), it is not surprising that they are a recently popular trait measured in seabirds (Supplementary Material Fig. A1), and across avian species and other taxa (Bonier et al. 2009; Busch and Hayward 2009). For example, Kitaysky and colleagues (Kitaysky et al. 1999, 2007, 2010; Satterthwaite et al. 2012) have demonstrated that metrics of CORT from Common Murres have the capacity to reflect current and past food abundance, as well as predict both a decrease in reproductive performance and the disappearance of individuals from a declining colony. Links between CORT and individual condition (Angelier et al. 2010) and reproductive success (Kitaysky et al. 2010) have also been reported in some seabird species, and while there is a general, but often complicated, assumption that variation in baseline GC levels should correlate with fitness (Bonier et al. 2009), empirical and meta-analytic studies have begun to indicate the often complex and context-dependent nature of the use of baseline GCs as direct links between environmental change and fitness (Lanctot et al. 2003; Hayward et al. 2011; Strasser and Heath 2013; Dantzer et al. 2014; Riechert et al. 2014; Madliger and Love 2016). As such, further work is needed to determine whether direct and predictive relationships exist between baseline GCs, fitness-related traits and environmental variation (Madliger and Love 2014; Thierry et al. 2014; Madliger and Love 2015). Crucially, while the frequent use of baseline CORT across seabird studies has resulted in a large body of research, we currently lack a quantitative assessment of how well these results predict key fitness-related traits that ultimately result in population-level changes.

The goal of the current paper was to evaluate whether variation in baseline plasma CORT can predict variation in metrics of fitness and the environment across seabird species. Such a study is both timely and useful for conservation and seabird biologists alike given that a large body of published research relating CORT levels to various metrics of fitness and environmental variation in seabirds exists with no general summary or consensus. We compiled all papers linking seabird baseline CORT to regularly measured traits related to seabird fitness: body condition, effort at foraging or nesting, and reproductive success, or to a key metric of environmental quality: food availability. Given that all of these traits are assumed to relate to individual success

(Ponchon et al. 2014) and are used in seabird literature to represent “success”, we consider these metrics ‘fitness-related traits’. Thus, to quantitatively assess the viability of using baseline CORT as a physiological biomarker, we used a phylogenetically controlled meta-analysis to analyse general trends between CORT and fitness-related traits using effect sizes. Given that it has commonly been assumed that elevated GCs have a negative association with fitness (although see Bonier et al. 2009) and low environmental quality (Kitaysky et al. 2010; Satterthwaite et al. 2012), we predicted a negative association between elevated baseline CORT and most fitness-related measures (e.g. body condition, food availability, reproductive success). However, given that baseline GCs are responsible for managing daily energetics (Angelier et al. 2007a), we also predicted a positive relationship with measures of individual effort in foraging and nesting.

Methods

We performed a systematic review and meta-analysis in which we collected data on the relationship between CORT and fitness-related traits from the literature and tested whether this relationship was influenced by the type of trait examined, study methodology, and the phase of breeding. We used Bayesian meta-regression techniques similar to Horvathova et al. (2011) and Prokop et al. (2012). We chose five fitness-related traits for the analysis based on their prevalence in the literature and the ease of obtaining similar data for both seabird researchers and conservation managers: body condition, foraging effort, nesting effort, food availability, and reproductive success (Table 1). We began with a Google Scholar and Web of Science search for the terms “corticosterone” and “seabird” (including papers available by June of 2015) across both correlative and manipulative studies that focused on links between baseline CORT and our chosen fitness-related traits (Fig. 2). We used only these general search terms to avoid missing papers that did not subscribe to the same terminology for the fitness-related traits that we chose. Manipulative studies included those, for example, where researchers either manipulated CORT levels directly (via implants or other supplementation), or manipulated the foraging ability or condition of individuals (via feather clipping or weight handicaps) and then examined impacts on CORT (e.g.

Table 1 Descriptions and biological relevance of fitness-related traits used for grouping of effect sizes in meta-analysis examining relationships between CORT and fitness in seabirds

Fitness-related trait	Specific traits compared to CORT	Biological relevance
Condition	Body mass, body condition index, age	Individual condition can be related to performance and fitness
Foraging effort	Number of foraging trips or dives, time at sea, effect of handicap	Altricial chicks and adult condition depend on parental foraging
Nesting effort	Attendance (time), chick feeding rate	Altricial chicks depend on parental foraging
Food availability	Sampled food abundance, fed vs. control individuals	Environmental trait, parents and chicks rely on food sources
Reproductive success	Number or percent of chicks hatched or fledged, clutch size, chick mass	Metrics of reproductive output should relate to population-level changes and individual fitness

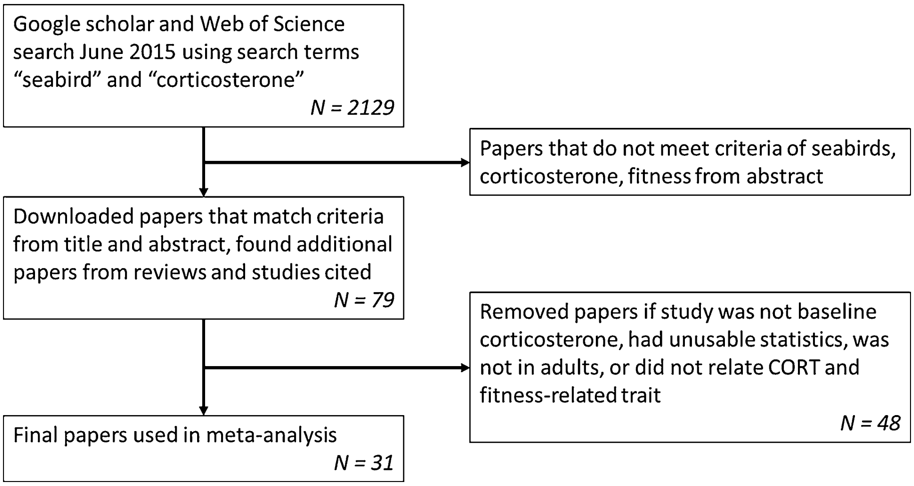


Fig. 2 Flow chart showing process of finding, selecting, and removing studies in meta-analysis

Angelier et al. 2007a; Leclaire et al. 2011). Our searches were limited to studies of adult seabirds (Orders Sphenisciformes (penguins), Procellariiformes (tubenoses), Pelecaniformes (gannets, cormorants, etc.), and the Charadriiformes (gulls and auks) during the breeding period (incubation, chick-rearing, or across both stages). Only studies of breeding adults were included due to prevalence in the literature (few studies of chick CORT and fitness) and due to the challenges of monitoring most seabirds outside of breeding. Of the 2129 papers returned from the initial search, the majority (2050) were rejected because they did not meet the basic criteria of pertaining to seabirds, CORT and fitness (Fig. 2). We retained 79 relevant papers for more thorough analysis by including all seabird papers studying CORT and traits that could be related to fitness. We then removed an additional 48 unusable studies from this group for one of more of the following reasons: (i) contained no direct examination of a relationship between CORT and a fitness-related trait; (ii) had unusable statistics or no statistics presented (i.e. only presented figures and/or estimates of *p* value); (iii) examined stress-induced as opposed to baseline CORT levels; or (iv) did not measure adults. Our final dataset contained 31 published papers (Supplementary Material Ref. A1) and one unpublished dataset (K. Elliott, G. Anderson and T. Gaston, unpublished data from collaborators) representing 16 species, with between one and 13 studies published on each species (Supplementary Material Table A1).

We examined all relevant papers and recorded each statistic, *p* value, and sample size presented for a reported relationship between CORT and a given fitness-related trait. Correlation coefficients (*r*) were used when available as they already represent effect sizes. In the absence of correlation coefficients, we used *F* values, *t* statistics, and Chi square values. Several studies presented none of this information, and as such *p* values were then used. We classified each effect size by: (i) type of study (correlational or experimental); (ii) reproductive stage (incubation, chick-rearing or across both stages); and (iii) fitness-related trait (adult condition, effort foraging, effort nesting, food availability or reproductive success, Table 1). These broad categories were used to determine whether certain fitness-related traits more strongly correlated with baseline CORT than others.

All analyses were performed in the R environment (3.1.2) using R Studio (0.99.879) as the working interface. Using the R package compute.es, we first calculated effect sizes as correlation coefficients (*r*) with a 95% confidence interval for every relationship found (Cooper et al. 2009). However, because *r* values are bounded at −1 and 1 and, therefore, do not meet the assumptions of parametric analyses, we transformed the values to Fisher's *Z* and calculated the appropriate sampling variance and standard error of variance according to the standard methods (Boncoraglio and Saino 2007). The Fisher's *Z* values and the corresponding sampling variance were used for further analysis.

Although all species in the current analysis are seabirds, and thus represent an already narrowed phylogenetic group, controlling for phylogenetic effects remains important (Adams 2008) as phylogenetic relationships could explain some variance in the relationship between CORT and fitness. To account for phylogenetic relationships in our analysis, we created a phylogenetic tree including all species in our dataset. First, we sampled 1000 trees based on our species list from birdtree.org, using the Hackett tree backbone (Hackett et al. 2008; Jetz et al. 2012). Then, we calculated the maximum clade credibility tree using tree annotater from the BEAST software, using the default settings (burnin = 0, node heights = median, output = maximum clade credibility tree) (Drummond et al. 2012).

The meta-analysis was conducted using the R package *MCMCglmm* version 2.22 (Hadfield 2010). We first created univariate models without intercepts with fitness-related trait, study method, and breeding stage independently as the fixed effect in each model, with phylogenetic and non-phylogenetic versions of each. These models are useful for comparing different levels of each variable to 0, but do not control for any covariates. We then created a series of eight multivariate models incorporating our three fixed effects and used the deviance information criteria (DIC) to choose the best models. A phylogenetic and non-phylogenetic version of each model was run, and species and individual study were included as random factors in all analyses. Our series of models (see descriptions in Table 2) tested first for an overall relationship between CORT and fitness (M1 and M2, non-phylogenetic and phylogenetic respectively), then for a relationship within each fitness-related trait (M3 and M4), then for fitness-related trait relationships by the study method (M5 and M6), and finally also including breeding stage as a fixed effect (M7 and M8).

For all models, we used the random effect prior of $V = 1$, $\nu = 0.02$, which specifies an inverse Gamma distribution with a degree of belief of 0.02, widely used in the statistical literature (Gelman and Hill 2007). Each MCMC chain was run for 1.1 million iterations, with a burn-in of 100,000, and a thinning interval of 1000, to produce a sample size (posterior distribution) of 10,000 samples. We checked convergence of model parameters using the potential scale reduction statistic (PSR; Gelman and Rubin 1992) and the multivariate potential scale reduction statistic (mvPSR; Brooks and Gelman 1998), calculated using the coda package (Plummer et al. 2006). To perform these tests, we ran two additional MCMC chains for each model (M1–M8), to produce a total of three chains per model (these additional chains were only used for convergence testing; all reported results are based on the first chain only). The potential scale reduction statistics compare among-chain and within-chain

Table 2 DIC values and variance components (study, species, and phylogeny) for eight models compared to determine whether fitness-related traits relate to baseline corticosterone in seabirds

Model	Fixed effects	DIC	Variance components		
			Study	Species	Phylogeny
M3	Fitness-related trait	150.022	0.176	0.166	–
M5	Fitness-related trait and study method	150.458	0.168	0.179	–
M4	Fitness-related trait	151.655	0.156	0.155	0.185
M6	Fitness-related trait and study method	152.007	0.144	0.164	0.205
M7	Fitness-related trait, study method, and breeding stage	155.339	0.167	0.161	–
M8	Fitness-related trait, study method, and breeding stage	156.709	0.148	0.153	0.199
M1	Intercept	157.758	0.238	0.166	–
M2	Intercept	159.452	0.217	0.150	0.174

All models included random effects for study and species and were run with and without phylogenetic control

variance, and should be less than 1.1 if chains are converging well (Gelman and Rubin 1992). In all our models, the PSR values were less than 1.01 for all parameters, and the mvPSR values were less than 1.02. Additionally, we visually checked chain mixing and autocorrelation using trace and density plots produced in coda. In all cases, chains appeared to be mixing well and had low autocorrelation.

Meta-analytical results can be influenced by a publication bias towards studies with large effect sizes. To test for signs of publication bias in our dataset, we conducted a visual and statistical analysis based on the inspection of funnel plots and running Egger's regression (Egger and Smith 1997), respectively. To account for phylogenetic relationships when assessing publication bias, we followed the modified approach to Egger's regression outlined in Nakagawa and Santos (2012), which uses the residuals from meta-regression models as the input variable in the Egger's regression to account for heterogeneity due to phylogeny and the random effect structure when assessing publication bias. We used our best fit model (Model 3) to conduct the Egger's regression analysis. When the intercept of the Egger's regression is significantly different from 0, the analysis can be interpreted as showing evidence of publication bias. In our case, $\beta_0 = -0.04$, 95% CI (-0.19, 0.11), $p = 0.57$. Additionally, we visually assessed funnel plots for evidence of asymmetry and outliers (Supplementary Material Fig. A2).

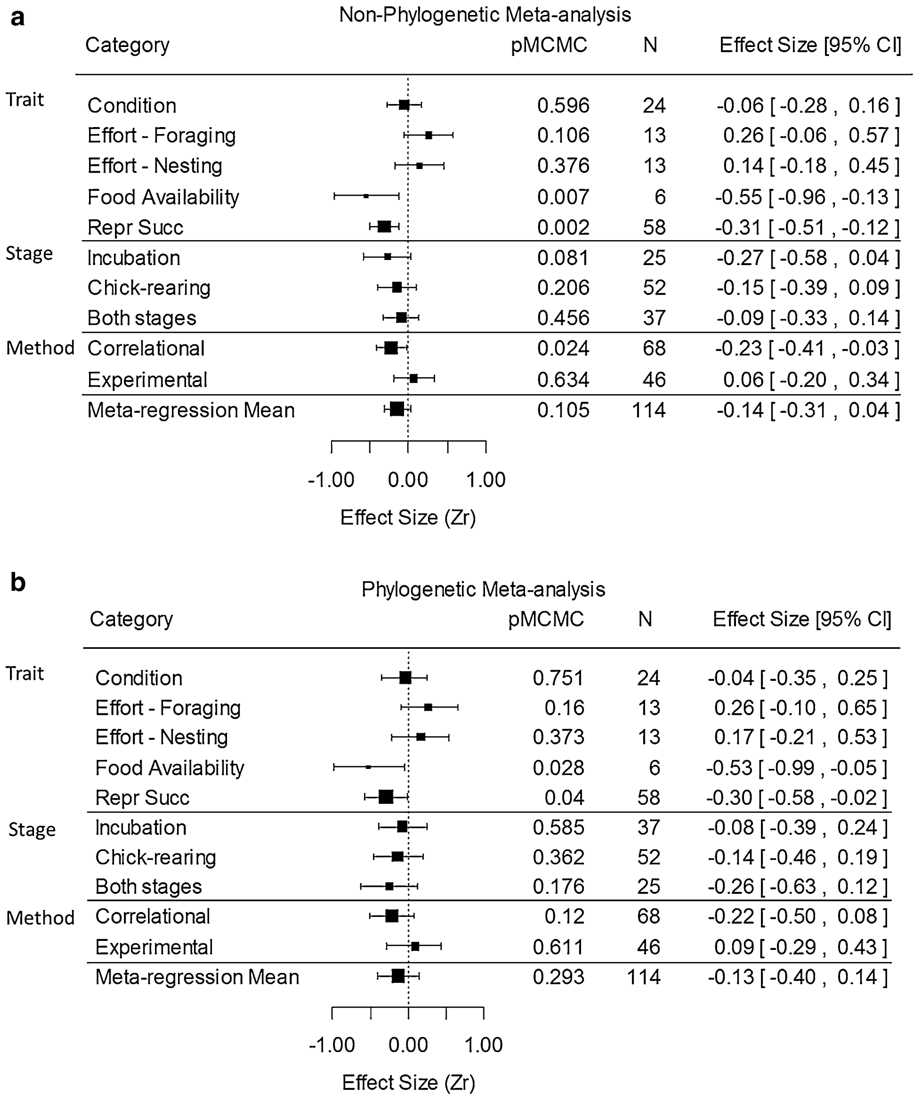
Results

In both the phylogenetic and non-phylogenetic univariate models, there were significant negative mean effect sizes for food availability [Fig. 3; non-phylogenetic meta-analytic mean $Zr = -0.56$, 95% CI (-0.96, -0.14)] and reproductive success [Fig. 3; non-phylogenetic meta-analytic mean $Zr = -0.31$, 95% CI (-0.50, -0.11)], but no strong correlation for any other fitness-related trait. There was no difference in correlation strength or direction between study method and breeding stage in the phylogenetic or non-phylogenetic models (Fig. 3). Overall, results indicate that baseline corticosterone was able to predict overall food availability and reproductive success, but not body condition, foraging effort or breeding effort (Fig. 3).

After comparing deviance information criterion (DIC) values between our eight models, M3–M6 were all within two DIC values and, therefore, were considered to be competitive models (Table 2). These four models all showed similar significant effects for food availability (Supplementary Material Table A2; Zr around -0.5) and reproductive success (Supplementary Material Table A2; Zr around -0.25). Models 5 and 6 both showed similar non-significant effects for the estimation of study method (Supplementary Material Table A2; experimental around 0.19). While no individual model could be classified as the strongest, the set of top models all suggested similar results.

Discussion

To proactively monitor the health of wildlife, researchers and conservation managers are increasingly looking towards sensitive individual-level metrics (e.g. physiology, behaviour, movement) in lieu of slower-responding demographic measures (Berger-Tal et al. 2011; Cooke et al. 2013, 2014; Madliger and Love 2015). Amongst the suite of physiological parameters proposed as biomarkers, GCs are often assumed to be reliable indicators of environmental impacts because they are closely tied to energetic management (Angelier et al. 2010) and food resources (Kitaysky et al. 2007) and, therefore, presumably to broader-scale environmental quality (Landys et al. 2006; Madliger et al. 2015). However, recent work has cautioned against the widespread



a Non-Phylogenetic Meta-analysis

	Category	pMCMC	N	Effect Size [95% CI]
Trait	Condition	0.596	24	-0.06 [-0.28 , 0.16]
	Effort - Foraging	0.106	13	0.26 [-0.06 , 0.57]
	Effort - Nesting	0.376	13	0.14 [-0.18 , 0.45]
	Food Availability	0.007	6	-0.55 [-0.96 , -0.13]
	Repr Succ	0.002	58	-0.31 [-0.51 , -0.12]
Stage	Incubation	0.081	25	-0.27 [-0.58 , 0.04]
	Chick-rearing	0.206	52	-0.15 [-0.39 , 0.09]
	Both stages	0.456	37	-0.09 [-0.33 , 0.14]
Method	Correlational	0.024	68	-0.23 [-0.41 , -0.03]
	Experimental	0.634	46	0.06 [-0.20 , 0.34]
	Meta-regression Mean	0.105	114	-0.14 [-0.31 , 0.04]

b Phylogenetic Meta-analysis

	Category	pMCMC	N	Effect Size [95% CI]
Trait	Condition	0.751	24	-0.04 [-0.35 , 0.25]
	Effort - Foraging	0.16	13	0.26 [-0.10 , 0.65]
	Effort - Nesting	0.373	13	0.17 [-0.21 , 0.53]
	Food Availability	0.028	6	-0.53 [-0.99 , -0.05]
	Repr Succ	0.04	58	-0.30 [-0.58 , -0.02]
Stage	Incubation	0.585	37	-0.08 [-0.39 , 0.24]
	Chick-rearing	0.362	52	-0.14 [-0.46 , 0.19]
	Both stages	0.176	25	-0.26 [-0.63 , 0.12]
Method	Correlational	0.12	68	-0.22 [-0.50 , 0.08]
	Experimental	0.611	46	0.09 [-0.29 , 0.43]
	Meta-regression Mean	0.293	114	-0.13 [-0.40 , 0.14]

Fig. 3 Mean effect sizes from four models run with fixed effects for trait, stage, method, and intercept only. Both non-phylogenetic **a** and phylogenetic models **b** are presented

use of GCs as biomarkers without first testing how well they relate to environmental variability and measures of fitness (Madliger and Love 2014, 2016). We performed a systematic review and meta-analysis to examine the correlation between baseline CORT levels and fitness-related traits in seabirds (recognized as important sentinel species due to their high trophic position and worldwide distribution; Piatt et al. 2007). The strength of relationships between baseline CORT and fitness varied based on the fitness trait. As predicted, food availability and reproductive success showed significant negative trends; however, body condition and effort at foraging and nesting showed no relationship. Moreover, our models including fitness trait as a fixed effect were much better supported than the null model, indicating that our fitness proxies were rooted in biological relevance and are, therefore, a practical means of grouping relationships between GCs and fitness in this taxonomic group.

Strength and biological relevance of CORT-fitness relationships

Body condition, which included measures such as body mass index, mass gain and fat level, did not generally correlate with baseline CORT. While individual studies did report significant results in these relationships (Supplementary Material Table A1), trends were generally weak. Although body condition can relate strongly to environmental conditions and to reproductive success (Wendeln and Becker 1999; Balbontín et al. 2012), individuals may be able to sacrifice individual condition to benefit offspring to varying degrees (Jacobs et al. 2013). Moreover, some Alcid species with extreme wing loading even exhibit adaptive mass loss (as a means to increase energetic efficiency) between the incubation and chick-rearing stages (Croll et al. 1991), potentially making relationships between body condition and fitness difficult to interpret. Variation in seabird life-history may also confound this relationship as seabird parents' ability to maintain their own condition during breeding may depend on the manner in which they carry food to offspring (whole or semi-digested prey) and the ability of offspring to survive for extended periods alone in the nest (i.e. need for constant feeding, predation risk, nest site dangers).

Neither effort during foraging, nor effort during nesting, correlated with baseline CORT, though both traits showed weak positive trends. Foraging effort included at-sea behaviours such as trip duration, amount of time diving for prey, and overall energy expenditure (i.e. daily energy expenditure—DEE). While these traits are predicted to relate to how well parents can care for their offspring and thus overall fitness (e.g. Weimerskirch et al. 2000), individual physiological management may obscure observable mean relationships with baseline CORT (Angelier et al. 2007b, c, 2009; Love et al. 2014). More specifically, it may be necessary to measure baseline CORT levels both before and after a foraging trip to fully appreciate how physiology may predict foraging or nesting success (Madliger and Love 2014; Love et al. 2014). This is likely true for nesting effort traits as well, given that time spent on the nest and chick feeding rate should relate to nesting success and fitness (e.g. Bukacinska et al. 2016).

In contrast to body condition and effort metrics, food availability showed a significant negative relationship with baseline CORT. Indeed, this was expected given baseline CORT's metabolic role and the fact that this pattern has been documented in other non-seabird avian species [e.g. barn swallow (*Hirundo rustica*): Jenni-Eiermann et al. 2008; white ibis (*Eudocimus albus*): Herring et al. 2011; zebra finch (*Taeniopygia guttata*): Lynn et al. 2010]. Importantly, food availability has been proposed as one of the major drivers of changes in seabird populations (Kitaysky et al. 2007). Despite this, we were only able to include six studies of food availability in our analysis, indicating that this is a potentially important area of research to expand upon in terms of pairing these data with the monitoring of circulating CORT levels (see below).

The strong relationship between baseline CORT and reproductive success could be similarly very important for proactively monitoring seabird populations and demographic trends. A review of the relationship between baseline GCs and reproductive success across multiple taxa (Bonier et al. 2009) reported mixed results (positive, negative and neutral), indicating that verification of the GC-fitness relationship is likely necessary in specific study systems (Bonier et al. 2009; Madliger and Love 2014). For example, one key consideration when testing relationships between GCs and fitness is life-history (Crossin et al. 2016a). Patterns in baseline GCs differ between life-history strategies, stages and even within reproductive stages and favour different aspects of trade-offs (e.g. Love et al. 2004; Williams et al. 2008). Our quantitative analysis of relationships between baseline CORT and reproductive success revealed that this trend is strongly negative across seabirds (a long-lived group of species) and across breeding stages. It is possible that because many long-lived seabird species are able to forgo reproductive attempts during harsher years to instead favour investment in future offspring, baseline CORT levels may more strongly correlate to reproductive outcomes, and may be more likely to show a negative relationship (Hau et al. 2010). For example, such species may be less likely to possess mechanisms to resist social and environmental stressors in favour of successful breeding, leading to correlations between high GC levels and reproductive abandonment (Wingfield and Sapolsky 2003). In general, in longer-lived seabird species, there is great potential for GCs to relate food availability to reproductive success, making them a strong biomarker if deployed within the appropriate system (e.g. Piatt and Harding 2007).

We were careful to recognize potential biases in the current literature. First, our meta-analysis as a whole showed no publication bias from unpublished significant studies or overly strong results in certain studies (based on Egger's regression). Second, we found no significant effect of the study method (experimental or correlational), though certain fitness traits are not represented evenly between methods. Most studies examining measures of foraging or nesting effort employed experimental methods to either increase the effort individuals faced (i.e. indirectly increase baseline CORT) or to increase baseline CORT directly. The lack of relationship with these effort-based traits and CORT, despite largely experimental methods, increases our confidence that CORT is not strong indicator of these traits overall. The opposite was true for the methodologies used in studies examining adult condition, food availability, and reproductive success where correlational methods characterized the majority of studies. Moving forward, additional studies employing experimental manipulations of CORT in the context of body condition, food resources, and reproductive outcomes will help to refine our appreciation of how baseline CORT may interact with the intrinsic and extrinsic environment to influence fitness in seabirds. While the focus in the current analysis was upon energetic management and hence variation in baseline corticosterone, we nonetheless appreciate that variation in the corticosterone response to environmental stressors has also been shown to be related to fitness metrics in seabirds (Kitaysky et al. 1999; Chastel et al. 2005; Angelier et al. 2015). As such, future work on this trait may indeed complement our findings here and may be useful in future predictions of how seabird populations and species will be expected to respond to further environmental change. Finally, we found no significant difference in mean effect size between correlational and experimental methods. Overall, we have confidence that the patterns we observed are not confounded by publication bias, or by a lack of consideration of methodological context.

Additional recommendations for using gcs as seabird biomarkers

It is clear from this meta-analysis and recent reviews (e.g. Bonier et al. 2009) that measurements of baseline GCs are not a perfect indicator for all fitness-related traits. Baseline GC levels are highly context-dependent because they can respond to both external environmental conditions and to internal changes in state; thus, fitness-related traits like condition and effort are likely to vary based on an individual's regulation of its physiology and its efficiency at gaining and using energetic resources (Busch and Hayward 2009; Madliger and Love 2015; Hennin et al. 2016). None of the studies completed in seabirds investigated how changes in baseline GCs over time (i.e. flexibility) related to fitness-related traits. However, there is growing evidence that the management of GCs over time may provide predictive capacity for fitness outcomes (Bonier et al. 2011; Ouyang et al. 2011, 2013; Love et al. 2014; Arlettaz et al. 2015). This may be particularly important to consider in seabird species that make extended foraging trips, requiring careful management of somatic resources while also tending to the offspring. We urge future studies to consider how changes in CORT, rather than simply static measures, may provide information on foraging success and ultimately reproductive outcomes.

The large amount of CORT research in diverse seabird taxa is testament to the desire of researchers and managers for a simple, holistic monitoring technique for this at-risk group of birds. The five fitness traits investigated in this meta-analysis showed varying relationships with baseline CORT, indicating that some traits may be more sensitive or more easily monitored with baseline GCs, while other traits may require more in-depth study to adequately assess their potential connection with baseline GCs. Indeed, seabird biologists have not sampled all fitness-related traits equally. For example, harder-to-measure traits such as food availability were measured in only four of the studies in our analysis, while body condition and reproductive success were measured in 16 and 17 studies, respectively. Focusing on certain environmental and fitness-related traits such as food availability and reproductive success is important, not simply because these relationships showed significance in our meta-analysis, but because these traits can best link GCs to environmental (food) changes and population demography. Studies measuring reproductive success in this meta-analysis looked at laying, hatching, and fledging. While some of these early fitness traits (e.g. clutch size) may be strongly linked to reproductive output in certain seabird species, focusing on fitness during chick-rearing should take precedence to best predict chick survival and recruitment to assess population changes.

One of the primary studies linking food availability to baseline CORT in black-legged kittiwakes (*Rissa tridactyla*) employed multiple methods of prey-species sampling as well as experimental methods to rigorously test the relationship between average (population-level) CORT levels and food availability in the local ocean environment (Kitaysky et al. 2010). Although time consuming and often expensive, further testing of this relationship should ideally follow these types of broad prey sampling methods and work towards increasing experimental methods to further examine causal links between food availability and changing GC levels. Furthermore, recent evidence suggests that unpredictable food availability (as opposed to simply the total amount of food available) may influence baseline GCs, indicating the importance of not only quantifying total energetic constraints, but also temporal and spatial fluctuations in availability to fully understand how individuals may respond to changing resource levels (Fokidis et al. 2012). If the relationship remains strong between CORT and food availability across additional studies/species, CORT levels in seabirds may indeed emerge as a very effective tool for sampling large areas of ocean productivity.

The next step to explicitly testing the usefulness of baseline CORT as a relevant biomarker in seabirds is the ability to scale up individual-level relationships to relate hormone levels to demographic patterns (adult/chick survival, breeding recruitment, population changes). If CORT predicts current year reproductive success, it will be worth testing how well it can predict future year success, though it will be important to appreciate possible time lags between baseline CORT and population changes, since baseline CORT levels in 1 year will likely relate to future population changes. Importantly, evidence is accumulating that elevated CORT levels may lead to carry-over effects, influencing patterns of migration, overwintering, and subsequent breeding. For example, Schultner et al. (2014) found that female black-legged kittiwakes exposed to a 3-day elevation of CORT during breeding left the breeding grounds earlier and spent a longer time at the wintering grounds than control females and CORT-treated males. In addition, Crossin et al. (2013, 2016b) found that feather CORT levels could successfully predict whether giant petrels (*Macronectes* sp.) and black-browed albatrosses (*Thalassarche melanophris*) would breed or defer breeding in the subsequent year. Paired with the experimental and correlational evidence in other avian and non-avian species indicating that GC measures can mediate carry-over effects (O'Connor et al. 2010; Mark and Rubenstein 2013; Sanderson et al. 2014), these patterns illustrate the potential for using GC measures to predict variation in future investment and success. Particularly in seabirds, which are comparatively long-lived, it will be important to determine whether CORT measures taken at different times may be able to forecast multi-year reproductive success across species (e.g. Angelier et al. 2010).

While baseline GCs do play an important role in connecting an individual's environment to its energetic demands, they are not the only mechanism that relate these states. Stress-induced or acute GC levels are also measured to relate to survival, though this relationship is also not fully understood, nor is it a clear pattern (Breuner et al. 2008). While stress-induced GCs have potential in providing this relationship, there are logistical challenges in measuring acute GCs, e.g. consistent restraint protocol and multiple blood sampling. Importantly, other metrics can be easier to collect (e.g. ease of field collection, avoiding stress-induced samples) and less expensive to measure analytically. For example, energetic metabolites (triglycerides, beta-hydroxybutyrate, and non-esterified fatty acids; Guglielmo et al. 2002) and measures of thyroid hormone can reflect the energetic state of individuals (Guglielmo et al. 2002; Elliott et al. 2013) and may additionally provide key information on foraging success and energetic constraints. Importantly, physiological panels that combine multiple energetic metrics may provide the best indication of overall current state and, therefore, future reproduction or survival potential, and are one way in which physiological measures are currently contributing to tangible conservation success (Madliger et al. 2016). Finally, behaviour is increasingly studied in seabirds due to the continued development of small GPS units that can accurately collect behaviour in flight and on the water (Gaston et al. 2013; Ponchon et al. 2014; Weimerskirch et al. 2014). By pairing physiological metrics with behaviour, researchers can examine how changes in physiology may manifest as alterations in reproductive behaviour (Cooke et al. 2014; Madliger and Love 2015), thereby better understanding the mechanisms by which environmental change may influence seabird populations. Overall, pairing multiple metrics of physiology, behaviour and environmental quality (e.g. food availability) will be paramount to assessing baseline GC levels as monitoring tools for seabird fitness and, ultimately, population demographics.

Author contribution statement GHS and OPL conceived and designed the search methodology. GHS searched the literature, collected the data, and calculated effect sizes. GHS and CJD analysed the data. GHS, CJD, and CLM wrote the manuscript with additional edits from OPL.

References

Adams DC (2008) Phylogenetic meta-analysis. Evolution 62:567–572. doi:10.1111/j.1558-5646.2007.00314.x

Angelier F, Clement-Chastel C, Gabrielsen GW, Chastel O (2007a) Corticosterone and time-activity budget: an experiment with Black-legged kittiwakes. Horm Behav 52:482–491. doi:10.1016/j.yhbeh.2007.07.003

Angelier F, Moe B, Clement-Chastel C et al (2007b) Corticosterone levels in relation to change of mate in Black-legged kittiwakes. Condor 109:668–674. doi:10.1242/jeb.089763

Angelier F, Shaffer SA, Weimerskirch H et al (2007c) Corticosterone and foraging behavior in a pelagic seabird. Physiol Biochem Zool 80:283–292. doi:10.1086/512585

Angelier F, Bost C-A, Giraudeau M et al (2008) Corticosterone and foraging behavior in a diving seabird: the Adelie penguin, *Pygoscelis adeliae*. Gen Comp Endocrinol 156:134–144. doi:10.1016/j.ygcen.2007.12.001

Angelier F, Giraudeau M, Bost C-AC-A et al (2009) Are stress hormone levels a good proxy of foraging success? An experiment with King Penguins, *Aptenodytes patagonicus*. J Exp Biol 212:2824–2829. doi:10.1242/jeb.027722

Angelier F, Wingfield JC, Weimerskirch H, Chastel O (2010) Hormonal correlates of individual quality in a long-lived bird: a test of the "corticosterone-fitness hypothesis". Biol Lett 6:846–849. doi:10.1098/rsbl.2010.0376

Angelier F, Wingfield JC, Parenteau C et al (2015) Does short-term fasting lead to stressed-out parents? A study of incubation commitment and the hormonal stress responses and recoveries in snow petrels. Horm Behav 67:28–37. doi:10.1016/j.yhbeh.2014.11.009

Arlettaz R, Nussle S, Baltic M et al (2015) Disturbance of wildlife by outdoor winter recreation: allostatic stress response and altered activity-energy budgets. Ecol Appl 25:1197–1212. doi:10.1890/14-1141.1.sm

Astheimer LB, Buttemer WA, Wingfield JC (1992) Interactions of corticosterone with feeding, activity and metabolism in Passerine birds. Ornis Scand 23:355–365. doi:10.2307/3676661

Balbontín J, Møller AP, Hermosell IG et al (2012) Lifetime individual plasticity in body condition of a migratory bird. Biol J Linn Soc 105:420–434. doi:10.1111/j.1095-8312.2011.01800.x

Benowitz-Fredericks ZM, Shultz MT, Kitaysky AS (2008) Stress hormones suggest opposite trends of food availability for planktivorous and piscivorous seabirds in 2 years. Deep Res Part II Top Stud Oceanogr 55:1868–1876. doi:10.1016/j.dsr2.2008.04.007

Berger-Tal O, Polak T, Oron A et al (2011) Integrating animal behavior and conservation biology: a conceptual framework. Behav Ecol 22:236–239. doi:10.1093/beheco/arq224

Bodey TW, Jessopp MJ, Votier SC et al (2014) Seabird movement reveals the ecological footprint of fishing vessels. Curr Biol 24:R514–R515. doi:10.1016/j.cub.2014.04.041

Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Funct Ecol 21:134–142. doi:10.1111/j.1365-2435.2006.01207.x

Bonier F, Martin PR, Moore IT, Wingfield JC (2009) Do baseline glucocorticoids predict fitness? Trends Ecol Evol 24:634–642. doi:10.1016/j.tree.2009.04.013

Bonier F, Moore IT, Robertson RJ (2011) The stress of parenthood? Increased glucocorticoids in birds with experimentally enlarged broods. Biol Lett 7:944–946. doi:10.1098/rsbl.2011.0391

Breuner CW, Patterson SH, Hahn TP (2008) In search of relationships between the acute adrenocortical response and fitness. Gen Comp Endocrinol 157:288–295. doi:10.1016/j.ygcen.2008.05.017

Brooks S, Gelman A (1998) General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 7:434–455. doi:10.2307/1390675

Bukacinska M, Bukacinski D, Spaans AL (2016) Attendance and diet in relation to breeding success in herring gulls (*Larus argentatus*). Auk 113:300–309

Busch DS, Hayward LS (2009) Stress in a conservation context: a discussion of glucocorticoid actions and how levels change with conservation-relevant variables. Biol Conserv 142:2844–2853. doi:10.1016/j.biocon.2009.08.013

Chastel O, Lacroix A, Weimerskirch H, Gabrielsen GW (2005) Modulation of prolactin but not corticosterone responses to stress in relation to parental effort in a long-lived bird. Horm Behav 47:459–466. doi:10.1016/j.yhbeh.2004.10.009

Chown SL, Gaston KJ (2015) Macrophysiology—progress and prospects. Funct Ecol 30:330–344. doi:10.1111/1365-2435.12510

Cooke SJ, O'Connor CM (2010) Making conservation physiology relevant to policy makers and conservation practitioners. Conserv Lett 3:159–166. doi:10.1111/j.1755-263X.2010.00109.x

Cooke SJ, Sack L, Franklin CE et al (2013) What is conservation physiology? Perspectives on an increasingly integrated and essential science. Conserv Physiol 1:1–23. doi:10.1093/conphys/cot001

Cooke SJ, Blumstein DT, Buchholz R et al (2014) Physiology, behavior, and conservation. Physiol Biochem Zool 87:1–14. doi:10.1086/671165

Cooper H, Hedges LV, Valentine JC (2009) Handbook of research synthesis and meta-analysis. Russell Sage Foundation, New York

Croll DA, Gaston AJ, Noble DG (1991) Adaptive loss of mass in thick-billed murres. Condor 93:496–502. doi:10.2307/1368181

Crossin GT, Phillips RA, Lattin CR et al (2013) Corticosterone mediated costs of reproduction link current to future breeding. Gen Comp Endocrinol 193:112–120. doi:10.1016/j.ygcen.2013.07.011

Crossin GT, Love OP, Cooke SJ, Williams TD (2016a) Glucocorticoid manipulations in free-living animals: considerations of dose delivery, life-history context and reproductive state. Funct Ecol 30:116–125. doi:10.1111/1365-2435.12482

Crossin GT, Phillips RA, Lattin CR et al (2016b) Physiological costs of reproduction and carryover effects in annually versus biennially breeding albatrosses (*Thalassarche* spp.). Antarct Sci

Croxall JP, Butchart SHM, Lascelles B et al (2012) Seabird conservation status, threats and priority actions: a global assessment. Bird Conserv Int 22:1–34. doi:10.1017/S0959270912000020

Dantzer B, Fletcher QE, Boonstra R, Sheriff MJ (2014) Measures of physiological stress: a transparent or opaque window into the status, management and conservation of species? Conserv Physiol 2:1–18. doi:10.1093/conphys/cou023

Descamps SS, Strøm H, Steen H et al (2013) Decline of an arctic top predator: synchrony in colony size fluctuations, risk of extinction and the subpolar gyre. Oecologia 173:1271–1282. doi:10.1007/s00442-013-2701-0

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973. doi:10.1093/molbev/mss075

Egger M, Smith GD (1997) Bias in meta-analysis detected by a simple, graphical test. Br Med J 315:629–634

Elliott KH, Welcker J, Gaston AJ et al (2013) Thyroid hormones correlate with resting metabolic rate, not daily energy expenditure, in two charadriiform seabirds. Biol Open 2:580–586. doi:10.1242/bio.20134358

Ewers RM, Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. Biol Rev Camb Philos Soc 81:117–142. doi:10.1017/S1464793105006949

Fokidis HB, Des Roziers MB, Sparr R et al (2012) Unpredictable food availability induces metabolic and hormonal changes independent of food intake in a sedentary songbird. J Exp Biol 215:2920–2930. doi:10.1242/jeb.071043

Frederiksen M, Mavor RA, Wanless S (2007) Seabirds as environmental indicators: the advantages of combining data sets. Mar Ecol Prog Ser 352:205–211. doi:10.3354/meps07071

Gaston AJ, Mallory ML, Gilchrist HG (2012) Populations and trends of Canadian Arctic seabirds. Polar Biol 35:1221–1232. doi:10.1007/s00300-012-1168-5

Gaston AJ, Elliott KH, Ropert-Coudert Y et al (2013) Modeling foraging range for breeding colonies of thick-billed murres *Uria lomvia* in the Eastern Canadian Arctic and potential overlap with industrial development. Biol Conserv 168:134–143. doi:10.1016/j.biocon.2013.09.018

Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge

Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472

Guglielmo CG, Hara PDO, Williams TD (2002) Extrinsic and intrinsic sources of variation in plasma lipid metabolites of free-living western sandpipers (*Calidris mauri*). Auk 119:437–445. doi:10.1642/0004-8038(2002)119[0437:EAISOV]2.0.CO;2

Hackett SJ, Kimball RT, Reddy S et al (2008) A phylogenomic study of birds reveals their evolutionary history. Science 320:1763–1768. doi:10.1126/science.1157704

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J Stat Softw 33:1–22. doi:10.1002/ana.22635

Halpern BS, Walbridge S, Selkoe KA et al (2008) A global map of human impact on marine ecosystems. Science 319:948–953. doi:10.1126/science.1149345

Harding AMA, Piatt JF, Schmutz JA (2007) Seabird behavior as an indicator of food supplies: sensitivity across the breeding season. Mar Ecol Prog Ser 352:269–274. doi:10.3354/meps07072

Hau M, Ricklefs RE, Wikelski M et al (2010) Corticosterone, testosterone and life-history strategies of birds. Proc R Soc B-Biol Sci 277:3203–3212. doi:10.1098/rspb.2010.0673

Hayward LS, Bowles AE, Ha JC, Wasser SK (2011) Impacts of acute and long-term vehicle exposure on physiology and reproductive success of the northern spotted owl. Ecosphere 2:1–20. doi:10.1890/ES10-00199.1

Hennin HL, Wells-Berlin AM, Love OP (2016) Baseline glucocorticoids are drivers of body mass gain in a diving seabird. Ecol Evol 6:1702–1711. doi:10.1002/ece3.1999

Herring G, Cook MI, Gawlik DE, Call EM (2011) Food availability is expressed through physiological stress indicators in nestling white ibis: a food supplementation experiment. Funct Ecol 25:682–690. doi:10.1111/j.1365-2435.2010.01792.x

Hoegh-Guldberg O, Bruno JF (2010) The impact of climate change on the world's marine ecosystems. Science 328:1523–1528. doi:10.1126/science.1189930

Horvathova T, Nakagawa S, Uller T (2011) Strategic female reproductive investment in response to male attractiveness in birds. Proc R Soc B Biol Sci 279:163–170. doi:10.1098/rspb.2011.0663

Hussey NE, Kessel ST, Aarestrup K et al (2015) Aquatic animal telemetry: a panoramic window into the underwater world. Science 348:1255642–1–1255642–10. doi:10.1126/science.1255642

Jacobs SR, Elliott KH, Gaston AJ (2013) Parents are a Drag: long-lived birds share the cost of increased foraging effort with their offspring, but males pass on more of the costs than females. PLoS One 8:e54594. doi:10.1371/journal.pone.0054594

Jenni-Eiermann S, Glaus E, Grüebler M et al (2008) Glucocorticoid response to food availability in breeding barn swallows (*Hirundo rustica*). Gen Comp Endocrinol 155:558–565. doi:10.1016/j.ygcen.2007.08.011

Jetz W, Thomas GH, Joy JB et al (2012) The global diversity of birds in space and time. Nature 491:444–448. doi:10.1038/nature11631

Kitaysky AS, Wingfield JCC, Piatt JFF (1999) Dynamics of food availability, body condition and physiological stress response in breeding Black-legged kittiwakes. Funct Ecol 13:577–584. doi:10.1046/j.1365-2435.1999.00352.x

Kitaysky AS, Piatt JF, Wingfield JC (2007) Stress hormones link food availability and population processes in seabirds. Mar Ecol Prog Ser 352:245–258. doi:10.3354/meps07074

Kitaysky AS, Piatt JF, Hatch SA et al (2010) Food availability and population processes: severity of nutritional stress during reproduction predicts survival of long-lived seabirds. Funct Ecol 24:625–637. doi:10.1111/j.1365-2435.2009.01679.x

Lanctot RB, Hatch SA, Gill VA, Eens M (2003) Are corticosterone levels a good indicator of food availability and reproductive performance in a kittiwake colony? Horm Behav 43:489–502. doi:10.1016/S0018-506X(03)00030-8

Landys MM, Ramenofsky M, Wingfield JC (2006) Actions of glucocorticoids at a seasonal baseline as compared to stress-related levels in the regulation of periodic life processes. Gen Comp Endocrinol 148:132–149. doi:10.1016/j.ygcen.2006.02.013

Leclaire S, Bourret V, Wagner RH et al (2011) Behavioral and physiological responses to male handicap in chick-rearing black-legged kittiwakes. Behav Ecol 22:1156–1165. doi:10.1093/beheco/arr149

Love OP, Breuner CW, Vézina F, Williams TD (2004) Mediation of a corticosterone-induced reproductive conflict. Horm Behav 46:59–65. doi:10.1016/j.yhbeh.2004.02.001

Love OP, Madliger CL, Bourgeon S et al (2014) Evidence for baseline glucocorticoids as mediators of reproductive investment in a wild bird. Gen Comp Endocrinol 199:65–69. doi:10.1016/j.ygcen.2014.01.001

Lynn SE, Stamplis TB, Barrington WT et al (2010) Food, stress, and reproduction: short-term fasting alters endocrine physiology and reproductive behavior in the zebra finch. Horm Behav 58:214–222. doi:10.1016/j.yhbeh.2010.03.015

Madliger CL, Love OP (2014) The need for a predictive, context-dependent approach to the application of stress hormones in conservation. Conserv Biol 28:283–287. doi:10.1111/cobi.12185

Madliger CL, Love OP (2015) The power of physiology in changing landscapes: considerations for the continued integration of conservation and physiology. Integr Comp Biol 55:1–9

Madliger CL, Love OP (2016) Do baseline glucocorticoids simultaneously represent fitness and environmental quality in an aerial insectivore? Oikos. doi:10.1111/oik.03354

Madliger CL, Semeniuk CAD, Harris CM, Love OP (2015) Assessing baseline stress physiology as an integrator of environmental quality in a wild avian population: implications for use as a conservation biomarker. Biol Conserv 192:409–417. doi:10.1016/j.biocon.2015.10.021

Madliger CL, Cooke SJ, Crespi EJ et al (2016) Success stories and emerging themes in conservation physiology. Conserv Physiol 4:1–17. doi:10.1093/conphys/cov057

Mark MM, Rubenstein DR (2013) Physiological costs and carry-over effects of avian interspecific brood parasitism influence reproductive tradeoffs. Horm Behav 63:717–722. doi:10.1016/j.yhbeh.2013.03.008

McEwen BS, Wingfield JC (2010) What is in a name? Integrating homeostasis, allostasis and stress. Horm Behav 57:105–111. doi:10.1016/j.yhbeh.2009.09.011

Moody AT, Hobson KA, Gaston AJ (2012) High-arctic seabird trophic variation revealed through long-term isotopic monitoring. J Ornithol 153:1067–1078. doi:10.1007/s10336-012-0836-0

Nakagawa S, Santos ESA (2012) Methodological issues and advances in biological meta-analysis. Evol Ecol 26:1253–1274. doi:10.1007/s10682-012-9555-5

O'Connor CM, Gilmour KM, Arlinghaus R et al (2010) Seasonal carryover effects following the administration of cortisol to a wild teleost fish. Physiol Biochem Zool 83:950–957. doi:10.1086/656286

Orians GH, Pearson NE (1979) On the theory of central place foraging. In: Horn DJ, Mitchell RD, Stairs GR (eds) Analysis of ecological systems. Ohio State University Press, Columbus, pp 154–177

Ouyang JQ, Sharp PJ, Dawson A et al (2011) Hormone levels predict individual differences in reproductive success in a passerine bird. Proc R Soc B Biol Sci 278:2537–2545. doi:10.1098/rspb.2010.2490

Ouyang JQ, Sharp P, Quetting M, Hau M (2013) Endocrine phenotype, reproductive success and survival in the great tit, *Parus major*. J Evol Biol 26:1988–1998. doi:10.1111/jeb.12202

Ozgul A, Childs DZ, Oli MK et al (2010) Coupled dynamics of body mass and population growth in response to environmental change. Nature 466:482–485. doi:10.1038/nature09210

Paleczny M, Hammill E, Karpouzi V, Pauly D (2015) Population trend of the world's monitored seabirds, 1950–2010. PLoS One 10:e0129342. doi:10.1371/journal.pone.0129342

Parsons M, Mitchell I, Butler A et al (2008) Seabirds as indicators of the marine environment. ICES J Mar Sci 65:1520–1526. doi:10.1093/icesjms/fsn155

Piatt JF, Harding AMA (2007) Population ecology of seabirds in cook inlet. In: Spies R (ed) Long-term ecological change in the Northern Gulf of Alaska. pp 335–352

Piatt JF, Sydeman WJ (2007) Theme Section: seabirds as indicators of marine ecosystems. Mar Ecol Prog Ser 352:199–309. doi:10.3354/meps07070

Piatt JF, Harding AM, Shultz M et al (2007) Seabirds as indicators of marine food supplies: cairns revisited. Mar Ecol Prog Ser 352:221–234. doi:10.3354/meps07078

Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11. doi:10.1159/000323281

Ponchon A, Grémillet D, Christensen-Dalsgaard S et al (2014) When things go wrong: intra-season dynamics of breeding failure in a seabird. Ecosphere 5:1–19. doi:10.1890/ES13-00233.1

Prokop ZM, Michalczyk Ł, Drobniak SM et al (2012) Meta-analysis suggests choosy females get sexy sons more than "Good genes". Evolution (N Y) 66:2665–2673. doi:10.1111/j.1558-5646.2012.01654.x

Ramos R, Garnier R, Gonzalez-Solis J, Boulinier T (2014) Long antibody persistence and transgenerational transfer of immunity in a long-lived vertebrate. Am Nat 184:764–776. doi:10.1086/678400

Ricklefs RE, WIlkelski M (2002) The physiology/life history nexus. Trends Ecol Evol 17:462–469

Riechert J, Becker PH, Chastel O (2014) Predicting reproductive success from hormone concentrations in the common tern (*Sterna hirundo*) while considering food abundance. Oecologia 176:715–727. doi:10.1007/s00442-014-3040-5

Sanderson JL, Young AJ, Hodge SJ et al (2014) Hormonal mediation of a carry-over effect in a wild cooperative mammal. Funct Ecol 28:1377–1386. doi:10.1111/1365-2435.12307

Satterthwaite WH, Kitaysky AS, Mangel M (2012) Linking climate variability, productivity and stress to demography in a long-lived seabird. Mar Ecol Prog Ser 454:221–235. doi:10.3354/meps09539

Schultner J, Moe B, Chastel O et al (2014) Corticosterone mediates carry-over effects between breeding and migration in the kittiwake *Rissa tridactyla*. Mar Ecol Prog Ser 496:125–133. doi:10.3354/meps10603

Strasser EH, Heath JA (2013) Reproductive failure of a human-tolerant species, the American kestrel, is associated with stress and human disturbance. J Appl Ecol 50:912–919. doi:10.1111/1365-2664.12103

Thierry A-M, Brajon S, Spee M et al (2014) Differential effects of increased corticosterone on behavior at the nest and reproductive output of chick-rearing Adelie penguins. Behav Ecol Sociobiol 68:721–732. doi:10.1007/s00265-014-1685-z

Weimerskirch H, Barbraud C, Lys P (2000) Sex differences in parental investment and chick growth in Wandering albatrosses: fitness consequences. Ecology 81:309–318

Weimerskirch H, Cherel Y, Delord K et al (2014) Lifetime foraging patterns of the wandering albatross: life on the move! J Exp Mar Bio Ecol 450:68–78. doi:10.1016/j.jembe.2013.10.021

Wendeln H, Becker PH (1999) Effects of parental quality and effort on the reproduction of common terns. J Anim Ecol 68:205–214. doi:10.1046/j.1365-2656.1999.00276.x

Williams CT, Kitaysky AS, Kettle AB, Buck CL (2008) Corticosterone levels of tufted puffins vary with breeding stage, body condition index, and reproductive performance. Gen Comp Endocrinol 158:29–35. doi:10.1016/j.ygcen.2008.04.018

Wingfield JC, Sapolsky RM (2003) Reproduction and resistance to stress: when and how. J Neuroendocrinol 15:711–724. doi:10.1046/j.1365-2826.2003.01033.x