Received: 21 September 2017 | Accepted: 24 October 2017
DOI: 10.1111/2041-210X.12934

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

A phylogenetically controlled meta-analysis of biologging device effects on birds: Deleterious effects and a call for more standardized reporting of study data

Thomas W. Bodey[1*] | Ian R. Cleasby[1*] | Fraser Bell[1] | Nicole Parr[1] | Anthony Schultz[2] | Stephen C. Votier[3] | Stuart Bearhop[1]

[1]Centre for Ecology and Conservation, University of Exeter, Penryn, Cornwall, UK

[2]GeneCology Research Centre, University of the Sunshine Coast, Sunshine Coast, Qld, Australia

[3]Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, UK

Correspondence
Thomas W. Bodey
Email: t.w.bodey@exeter.ac.uk
and
Ian R. Cleasby
Email: i.r.cleasby@gmail.com
and
Stuart Bearhop
Email: s.bearhop@exeter.ac.uk

Funding information
ERC grant STATEMIG, Grant/Award Number: 310820

Handling Editor: Emmanuel Paradis

Abstract

1. The use of biologging devices continues to increase with technological advances yielding remarkable ecological insights and generating new research questions. However, as devices develop and are deployed more widely, there is a need to update our knowledge of the potential ethical impacts to allow scientists to balance these against the knowledge gained.
2. We employed a suite of phylogenetically controlled meta-analyses on a dataset comprising more than 450 published effect sizes across 214 different studies to examine the effects of biologger tagging on five key traits in birds.
3. Overall, we found small but significant negative effects of tagging on survival, reproduction, parental care. In addition, tagging was positively associated with foraging trip duration, but had no effect on body mass. Meta-regressions revealed that flying style, migration distance and proportional tag mass were significant influences producing these deleterious effects, with attachment type and position additionally important covariates influencing survival- and reproduction-based effect sizes.
4. There was a positive correlation between the effects of tagging on survival and reproduction, highlighting that effects may be cumulative, with the full effects of tagging not necessarily apparent in studies focused on single traits. We discuss the tradeoff between these negative effects and the advances gained through the use of biologgers.
5. Finally, given the number of studies from our initial literature search that lacked sufficient data for inclusion in analyses, we provide recommendations on the essential information that all biologging studies should report in order to facilitate future assessments of impacts on animals.

KEYWORDS
Avian, geolocator, GPS, mass, radio, satellite, tag, tracking

*These authors contributed equally to this work.

 © 2017 The Authors. Methods in Ecology and Evolution © 2017 British Ecological Society

1 | INTRODUCTION

The biologging revolution, with the use of small, lightweight devices to record spatial and physiological parameters of animals, has increased our understanding of the natural world (Hussey et al., 2015; Kays, Crofoot, Jetz, & Wikelski, 2015). It has transformed the study of migration (Milner-Gulland, Fryxell, & Sinclair, 2011), and has provided remarkable insights into foraging (Adachi et al., 2016; Bodey et al., 2014; Wakefield et al., 2013) and physiology (Bishop et al., 2015; Watanabe, Goldman, Caselle, Chapman, & Papastamatiou, 2015), generating a new understanding of these processes in many animal species. Increasingly, biologging devices carried by animals are being used as monitoring tools, providing insights into abiotic environmental processes, ecosystem function and human activities (Kays et al., 2015).

While there is a natural temptation to exploit this technology in order to reveal hidden processes and gain deeper insights, there are important questions surrounding the ethics of obtaining such data, and the reliability of estimates derived from it. An overly instrumented animal may behave in atypical ways, for example, through reduced movement, excessive comfort behaviours and, in the most extreme cases, death (Thaxter et al., 2016). This is of particular relevance as new and more sophisticated devices become available that are extremely attractive due to an increase in quantity/quality of data recorded and thus the questions that can be addressed (Hussey et al., 2015; Kays et al., 2015). Obtaining accurate and “typical” data is key to all scientific enquiry and there is a long-standing concern with instrumenting animals and the effects this may have (Casper, 2009; Kenward, 2001; Vandenabeele, Shepard, Grogan, & Wilson, 2012; Wilson & McMahon, 2006). Indeed, guidelines on animal welfare in research have evolved significantly in many countries over the period that biologger use has increased (Anonymous 2012; Casper, 2009; Sergio et al., 2015; Wilson & McMahon, 2006). Such concerns often play a critical role in the choice of biologging device and the attachment type a particular study will adopt. While the literature can offer guidance, protocols are frequently copied from “similar” species with the expectation that what works for one will apply to another (Casper, 2009; Kenward, 2001; Vandenabeele et al., 2012; Wilson & McMahon, 2006). Assessing the extent to which this is a valid approach is made difficult due to the range and type of data reported as a result of different study aims, the disparate publishing locations of the research, and the reduced likelihood of publication of negative consequences (but see: Thaxter et al., 2016). In addition, although reviews and “how to” guides on tagging and animal welfare have previously sought to address this issue (Casper, 2009; Phillips, Silk, Croxall, Afanasyev, & Briggs, 2004; Vandenabeele et al., 2012; Wilson & McMahon, 2006), some rules-of-thumb, such as the 5% or 3% proportion of body mass that devices may not exceed, have become accepted “standards” despite limited evidence as to their broad applicability (Barron, Brawn, & Weatherhead, 2010; Casper, 2009; Gannon & Sikes, 2007; Kenward, 2001; Phillips et al., 2004).

In birds, there have been concerns as to the effects of biologging devices on the behaviour, and particularly the survival, of instrumented individuals for over a decade (Casper, 2009; Phillips et al., 2004; Sergio et al., 2015). The incorrect device deployment can affect survival rates, and even carefully considered choices may reduce return rates in migrating birds (Thaxter et al., 2016), or impair foraging or breeding success (Barron et al., 2010). In efforts to summarize the impacts of tagging, negative effects have generally been found in meta-analyses looking at particular tag types (e.g. geolocators Costantini & Møller, 2013), particular species groups (e.g. waders Weiser et al., 2016) or across the literature as a whole (Barron et al., 2010). However, as new empirical evidence accumulates, meta-analyses need to be revised, especially as effect sizes (which provide a measure of the magnitude of a treatment effect) tend to decline over time in fields such as ecology (Jennions & Møller, 2001), and more powerful meta-analytical techniques develop (Hadfield & Nakagawa, 2010). In addition, the available technology is continually changing and, in particular, miniaturizing, resulting in an increasing range of positional and physiological devices being deployed across an expanding range of species (Figure S1). There is, therefore, an urgent need for a more complete assessment of the effects of biologging devices in order to provide researchers with the best information and guidance. Moreover, previous studies have focused on limited sets of traits or tag types, or combined data from wild and captive organisms operating under very different constraints, and none have controlled for phylogeny.

Here, we provide the first phylogenetically controlled meta-analysis of the principle biologging devices (listed in methods) and the extent of their impacts on birds for a suite of key traits: survival, reproduction, body mass, parental care and foraging behaviour. We use meta-regression techniques (Nakagawa & Santos, 2012) to explore how factors such as device mass, attachment method or position moderate the effect of tagging. We also employ a multivariate procedure to examine whether the effects of tagging are correlated between traits across species, providing a more complete assessment of potential cumulative impacts. Together this enables us to highlight the (in)appropriateness of various attachment types for different species groups, particularly if these are considered functionally in terms of flying style, migration distance, etc. rather than purely taxonomically. Lastly, we provide specific recommendations as to the type of data that should be supplied with all publications that employ biologgers in order that the effect of current, and new, technologies can be thoroughly assessed through similar approaches. This will ensure that both animal welfare and data quality standards are as high as possible, and that recommendations for best practice can be re-evaluated and updated as necessary.

2 | MATERIALS AND METHODS

2.1 | Literature search and data compilation

We searched the literature using ISI Web of Science (http://wok.mimas.ac.uk/) and Google Scholar (https://scholar.google.co.uk/), using the search term *bird** alongside device-specific terms: *ARGOS*, *Geolocator*, *GLS*, *GPS*, *GSM*, *PTT*, *Satellite*, *UHF* and *VHF*. As the latter two terms produced few results, we combined *radio and bird** with additional terms *backpack*, *collar*, *harness*, *satellite*, *telemetry* and *transmitter*. Using this approach for other device-specific terms did not produce additional relevant publications. Searches were conducted for papers published during the period January 2009 to April 2016. We incorporated papers published pre-2009 by adding all studies cited within reviews of specific tag types or species groups (Barron et al., 2010; Costantini & Møller, 2013; Godfrey & Bryant, 2003; Phillips et al., 2004; Vandenabeele et al., 2012; Watanabe & Nathan, 2016; Weiser et al., 2016). We also performed a forward search to include any studies citing these reviews. Unpublished datasets were not solicited to reduce the risk of biasing effect size estimates (Jennions, Lortie, Rosenberg, & Rothstein, 2013).

2.2 | Inclusion/exclusion criteria

Following the literature search, studies were retained based upon the following criteria: (1) only studies conducted in the wild were included, with studies of captive or released individuals excluded, (2) The sample size for the study must have ≥5 individuals, (3) Data must be provided on both tagged and control birds; studies referencing earlier data or other studies as a control were excluded. Control birds were not fitted with biologger tags, but were fitted with identification markers (e.g. colour rings) and are best considered as procedural controls having been caught and handled, (4) A suitable effect size estimate, or sufficient information for an effect size calculation must be provided. These criteria ensured that studies provided suitable information for undertaking a meta-analysis. Although our initial collation from search terms totaled over 13,000 publications, deploying these criteria ultimately produced 451 effect sizes across 214 different studies. A detailed breakdown of our data search is provided in the Appendix (Figure S1).

2.3 | Data extraction

Effect sizes came from numerous sources: (1) Direct reports e.g. Cohen's *d* (2) measures with magnitude and direction e.g. regression coefficients (3) raw numbers e.g. contingency tables and (4) inferential test statistics e.g. t or χ^2 values. Data were converted into Fisher transformed correlation coefficients *Zr*, following standard formulae (Nakagawa & Cuthill, 2007) and *Zr* was used as the response variable in our meta-analyses. Values of Zr = 0.1, 0.31 and 0.55 are used to refer to small, medium and large effects respectively. We used *Zr* as our effect size metric because it is normally distributed, permitting parametric analysis, and the sampling variance of *Zr* is easily estimated as: $1/(N-3)$, where N represents the sample size on which the given effect size is based (Lipsey & Wilson, 2001).

2.4 | Potential impacts on vital rates and other traits

In order to assess the impact of devices on birds, we categorized effect sizes based on the nature of the hypothesis tested and data available. The key traits we examined for tagging effects were as follows: survival, reproduction, body mass, parental care and foraging behaviour (Table S1). We calculated separate effect sizes for each measure from every independent population tested within a study (for totals see Table 1 and Figure S1). Thus, if tag effects were examined across two distinct populations in a given study, two independent effect sizes were calculated. Similarly, if studies assessed tag effects on independent samples of individuals over multiple years, separate effect sizes were calculated for each study year (to account for potential non-independence when a study provides more than on effect size we included study ID as a random effect, see below). We also collected data on a number of potentially important moderator variables including attachment location, flying style, etc. to allow us to examine what factors influence the effect of tagging (Table 2).

2.5 | Phylogenetic meta-analysis

All analyses were conducted in the R environment (R Development Core Team, 2016), using the R package MCMCglmm (Hadfield, 2010) unless specified otherwise. We ran separate Bayesian multilevel meta-analyses for each of the five key traits. For each category, we present the results of an intercept-only model (Model A), providing the overall mean effect size across studies (i.e. standard meta-analysis), with random effects for: study, species, year and phylogeny, and the measurement error variance for each individual effect size set as $1/(N-3)$. The inclusion of a phylogenetic random effect allowed us to account for the potential non-independence of data due to shared ancestry. To incorporate phylogeny, we used the Ericson backbone tree from Jetz, Thomas, Joy, Hartmann, and Mooers (2012), downloaded from: http://birdtree.org/. As an index of the magnitude of the phylogenetic signal in the data, we calculated the phylogenetic heritability H^2 (Hadfield & Nakagawa, 2010).

For each key trait category, we then ran single factor meta-regression models (Model B) with the same random effects listed above, and entering each moderator variable in turn to obtain parameter estimates for each level in each factor. The continuous variables proportional tag mass and duration of deployment were included to control for their potential effect, but were removed if model Deviance Information Criterion (DIC) scores indicated this did not reduce model fit. In addition, we tested whether there was a change in

TABLE 1 Results from the random-effects meta-analyses on the effect of tagging across the five different key traits. Significant results are highlighted in bold. *N* is the number of effect sizes/individual studies included for the relevant variable

Variable	Mean effect size (*Zr*)	Lower 95% CRI	Upper 95% CRI	*N* effect sizes/studies
Survival	**−0.064**	**−0.111**	**−0.019**	140/103
Reproduction	**−0.051**	**−0.10**	**−0.005**	149/77
Body mass	−0.004	−0.094	0.094	75/62
Parental care	**−0.13**	**−0.24**	**−0.025**	50/27
Foraging duration	**0.13**	**0.004**	**0.264**	37/28

TABLE 2 Description of the moderator variables extracted during the literature search for consideration in meta-regressions models and details on how data on each was collected

Data	Description and extraction rules
Organism variables	
Body mass (g)	We used the mean body mass recorded in the study or collated from the data if available. If this was not reported, we used the species body mass from Dunning (2008). If species were sexually dimorphic in size and study comparisons were reported separately, we used this separation. If comparisons were merged, we took a mean across the masses of the sexes
Sex	Only males tagged, only females tagged or both sexes tagged
Age	Only adults tagged, only juveniles tagged or both adults and juveniles tagged
Flying style	Flight mode of species studied. Defined as follows: Flapping, Soaring, both flapping and soaring, or flightless Watanabe and Nathan (2016)
Migration strategy	Birds classed as long-distance migrants, short-distance migrants, partial migrants or non-migrants BirdLife International (2016)
Device variables	
Device mass	If study comparisons were separated by mass then we used this separation. If comparisons merged all data, then we took the mean device mass unless studies had <10% of a heavier device type, where we took the modal device mass
Device mass as % of body mass	This was calculated based on the combined or separated data as detailed above
Device mass > 1% of body mass	Was device mass >1% of species body mass? Initially, tag weights were split into three categories: (A) Tags <1% of the species body mass, (B) Tags between 1% and 3% of species body mass, and (C) Tags >3% of species body mass. These values are common benchmarks cited in the tagging literature. However, we never found differences in the estimated effect size between categories B and C and so these were amalgamated into one category representing tags that were >1% of species body mass
Length of deployment	We used the mean time length recorded in the study, or extracted from summary tables of individuals if available
Statistical data	
How tested	What statistical methods were used to test tag effects, e.g. Contingency table, means and standard deviations, etc
Within or among individual?	Data collected for body mass effect sizes only. Was effect of body mass assessed comparing across individuals or within individuals (change in mass)

the magnitude of reported effect sizes over time. However, we found no time trend in reported effect sizes in any meta-analysis model and this term was subsequently removed from all models (for more details see Appendix). We present I^2 values as a measure of inconsistency in effect sizes across studies. Here, I^2 is generally defined as the ratio of true heterogeneity to the total variance across studies (Borenstein, Hedges, Higgins, & Rothstein, 2009), with I^2 benchmark values of 25%, 50% and 75% representing low, moderate and high values respectively (Higgins, Thompson, Deeks, & Altman, 2003). Due to the inclusion of random effects, we calculated a modified version of I^2 following Nakagawa and Santos (2012). For meta-regression models, the calculation of I^2 is inappropriate, so instead we report the marginal and conditional R^2 values for each meta-regression model following Nakagawa and Schielzeth (2013).

2.6 | Multivariate meta-analysis

We performed a multivariate meta-analysis (Model C), using weighted species mean effect sizes as the unit of analysis, using the WinBUGS program (Lunn, Thomas, Best, & Spiegelhalter, 2000; Cleasby & Nakagawa, 2012). This allowed us to estimate whether the effects of tagging on key traits are correlated across species. Due to the nature of reported data we often encountered missing values whenever a species provided an effect size for one category but not another. To account for this, we used Bayesian data augmentation to prevent bias when estimating correlations among effect size categories (for more details see Appendix, for WinBUGS code see Cleasby & Nakagawa, 2012).

For all meta-analysis models, we used parameter-expanded priors for the random effects, running 3 MCMC chains for 500,000 iterations, with a thinning interval of 25 after a burn-in of 100,000. Autocorrelation between posterior samples was <.1 for all estimated parameters, and the Gelman-Rubin diagnostic was <1.2 for all parameters, indicating chain convergence. Results are reported with 95% Bayesian credible intervals (CRIs), and also more conservative 80% CRIs due to the importance of avoiding negative impacts on study organisms.

2.7 | Assessing publication bias

To test for possible publication bias, we ran Egger's regression test (Egger, Davey Smith, Schneider, & Minder, 1997) on the residuals from our meta-analysis models (Nakagawa & Santos, 2012). In addition, we ran trim-and-fill analyses (Duval & Tweedie, 2000) on model residuals to identify and correct for potentially missing studies (more details in Appendix).

3 | RESULTS

3.1 | Model A: Standard random effects meta-analysis

On average, tagging birds produced small but significant impacts on four of the five key traits examined, with tagged birds suffering reductions in survival, reproductive success and parental care (Table 1,

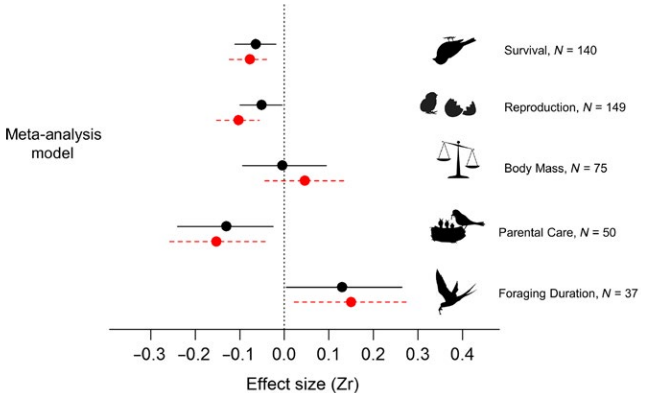

FIGURE 1 Forest plot of the meta-analytic means looking at the effect of tagging across each of the meta-analyses conducted. Black points and lines represent estimates from the standard random effects meta-analysis. Red points and dashed lines represent effects sizes after correcting for potential missing studies identified by trim-and-fill tests

Figure 1), and with increases in foraging trip length (Table 1), compared with controls. There was no evidence of significant publication bias in any of these analyses (Table S1). However, trim-and-fill methods suggested that there were potential missing studies, and adjusting for these slightly increased the negative effect of tagging across all traits (Table S1, Figure 1). There was no evidence that tagging influenced the body mass of individuals, even after adjusting for potential missing studies, and again no evidence of publication bias.

Across all five meta-analyses, the random effects included explained little of the variation in effect sizes, and there was no evidence of phylogenetic heritability in tag effects (Table S2). There was some evidence of between-species variation in survival, and between-study variation in body mass, although the variance component intervals were quite wide (Table S2). I^2 values indicate that there was high heterogeneity in survival and reproduction effect sizes, but lower levels of heterogeneity in relation to body mass, parental care and foraging trip length (Table S1).

3.2 | Model B: Meta-regression on key traits with moderator variables

3.2.1 | Survival

Effect sizes varied across methods of tag attachment, with harness and tailmount attachment associated with significant negative effects on survival at 95% CRI, while leg band and poncho methods also produced negative effects at more conservative 80% CRI (Table S3, Figure 2a). Effect sizes were non-significant for other attachment methods, however in certain cases, available sample sizes were small, which will reduce statistical power. Reductions in survival when tags were placed on a bird's back or tail are probably reflective of the negative effect of harnesses and tail-mounts.

Examination of flight type revealed that tagged birds with flapping flight experienced reduced survival compared to other styles. Negative effects were also greatest in species with long migration distances (Table S3). Tagging was also associated with lower survival in studies in which only one sex was tagged, but was not associated with survival in studies in which only juveniles were instrumented (though it should be noted, the effect on juveniles was based on a sample size of 10). Neither proportional tag mass (β = −0.0051; 95% CRI: −0.039 to 0.031) or deployment duration (β = 0.017; 95% CRI: −0.006 to 0.039) significantly influenced survival. However, when categorizing proportional tag mass as above or below 1% of species' body mass, we found a negative effect of tagging upon survival when tags were >1%, but no effect when tags were <1% of body mass (Figure 2a). The conditional R^2 (variance explained by both fixed and random factors) of our survival meta-regression models was high (R^2_{COND} = .91). However, the marginal R^2 (variance explained by fixed factors only) was lower (R^2_{MARG} = .18).

3.2.2 | Reproduction

The significant effects of tags on reproduction were largely limited to those associated with tagging with neck collars (negative impacts of collars, neck as a tag position and flapping flight at 95% CRI, Table S4, Figure 2b). Although, for certain tagging methods our sample size was low, which may explain non-significant results in some instances. As with survival, the effects of tagging on reproduction appeared more negative in birds with flapping flight compared to other flight styles, with long-distance migrants also negatively affected at the 80% CRI (Table S4). Tags also had negative impacts on reproduction when >1% of species' body mass (Figure 2b). There was no evidence that deployment duration (β = 0.007; 95% CRI: −0.036 to 0.051) influenced reproduction effect sizes. As with survival the fixed factors explained little of the variation in effect sizes (R^2_{COND} = .51; R^2_{MARG} = .14).

3.2.3 | Body mass

None of the moderator variables assessed was significant at the 95% CRI (Table S5). However, 80% CRIs suggest a trend for collars and neck attachments (collars and ponchos) to reduce body mass (Table S5, although estimates were based on small sample sizes). Similarly, at 80% CRIs, studies tagging only one sex showed greater negative effects than those tagging both sexes. Neither deployment duration, proportional or categorical tag mass produced significant effects (Table S5). The R^2 values were: R^2_{COND} = .92 and R^2_{MARG} = .31.

3.2.4 | Parental care

Effect sizes for parental care were positively associated with tail-mount attachment methods (Table S6), although this result was based on a very small sample size (two effect sizes from two studies) so should be treated with caution. Otherwise, 80% CRIs revealed parental care tended to decline when tags were fitted internally or via adhesive, and when attached on the back (Table S6). No other attachment methods were associated with parental care effect sizes, and there was no evidence for an association with proportional tag mass or deployment duration, although there was a tendency for tags weighing >1% body mass to reduce parental

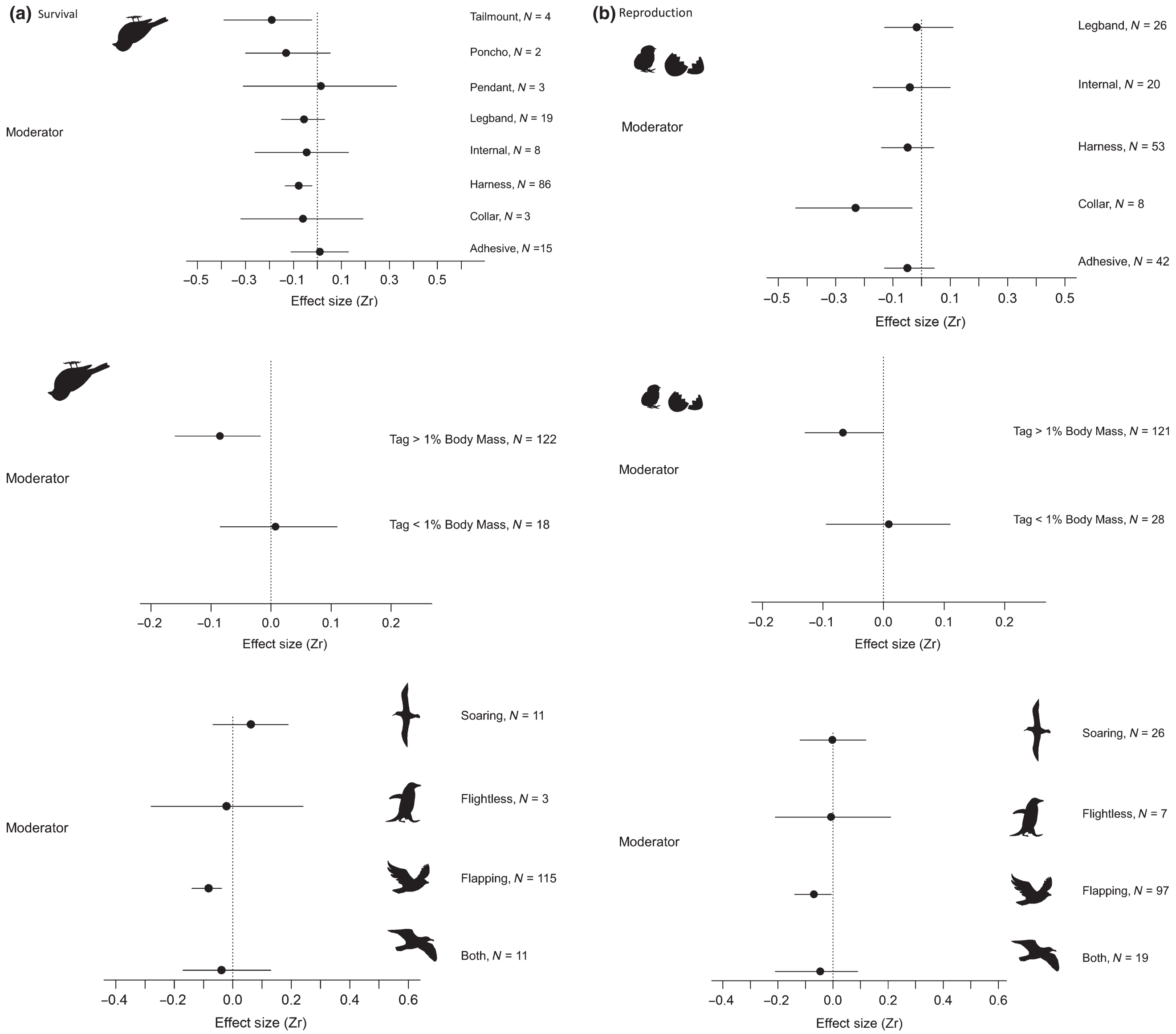

FIGURE 2 Forest plots showing the effect of key moderator variables—tag attachment method, whether a tag weighed >1% of body mass and flying style on survival (a) and reproduction (b). N = number of effect sizes

care (Table S6). The R^2 values were as follows: R^2_{COND} = .88 and R^2_{MARG} = .39.

3.2.5 | Foraging trip duration

Fitting tags to male birds significantly increased foraging trip durations (Table S7). Otherwise there were tendencies at 80% CRIs for birds to increase foraging trip duration when tags were placed on the back, and for all flying styles except soaring (Table S7). Deployment duration, proportional or categorical tag mass were not associated with foraging effect sizes. The R^2 values were: R^2_{COND} = .94 and R^2_{MARG} = .11.

3.3 | Model C: Multivariate meta-analysis

Effect sizes estimated using a multivariate meta-analysis were similar to those estimated by standard meta-analysis models (model A). Thus, results are qualitatively unchanged whether analysed at the population or species level (Table S8). There was a positive correlation between survival and reproduction effect sizes at the species level (Figure 3), indicating that when tagging negatively affects a species' survival rates, it is also likely to reduce reproduction and vice versa. There was little evidence of correlations between other effect size categories, although relatively small sample sizes limit the precision of correlation estimates (Table S9).

4 | DISCUSSION

Overall, phylogenetically corrected meta-analyses revealed that tagging birds had small, but significant negative effects on a number of key traits. This was confirmed by both standard and multivariate meta-analyses, with tagging associated with small reductions in survival,

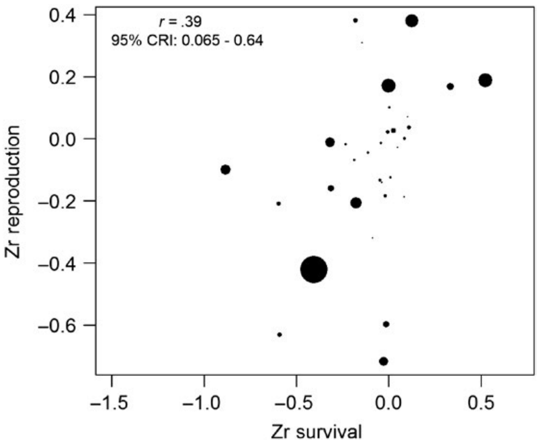

FIGURE 3 Correlation between survival-based and reproduction-based effect sizes estimated at the species-level via multivariate meta-analysis regression. Data points on plot are proportional to sample size per species (*N* survival effect size estimate + *N* reproduction effect size estimate)

reproductive success and parental care, and an increase in foraging trip durations (Figure 1). The only key trait unaffected by tagging was individual body mass. Moderator variables that repeatedly influenced key trait effect sizes at both 95%, and more conservative 80%, CRIs were flying style, migration length and the weight of tag relative to body mass. In addition, attachment methods and device position were associated with effects on survival and reproduction. We found little evidence for a phylogenetic signal, suggesting that these conclusions are not simply a result of more closely-related species suffering similar impacts. There was also no evidence found of publication bias, and accounting for likely missing studies led to a slight strengthening of all effect sizes.

Standard meta-analysis models also identified a large degree of heterogeneity in both survival and reproduction. While high heterogeneity is expected in most ecological studies, our estimates of I^2 for survival and reproduction are similar to the median I^2 of 85% reported across ecology (Senior et al., 2016), and significant heterogeneity was reported by Costantini and Møller (2013) in their meta-analysis of geolocator effects. Heterogeneity was much lower in the case of body mass, parental care and foraging trip duration. In particular, the low I^2 seen in the meta-analysis of parental care suggest declines in parental care may reflect a conserved response across species (Senior et al., 2016); possibly as a means by which parents shift some of the costs of tagging onto their offspring (Mauck & Grubb, 1995).

In many studies, the ratio of tag mass to body mass is frequently used to justify tag choice, taking precedence over other considerations such as tag shape or profile (Bowlin et al., 2010; Vandenabeele et al., 2012). Most studies adhere to either a 5% or 3% "rule" when fitting devices. However, we found that negative effects upon survival, reproduction and parental care were apparent only when tags weighed more than 1% of species' body mass. Thus, our results again do not support a "5% rule" (Barron et al., 2010; Kenward, 2001), but also provide little evidence to support a "3% rule" either (Phillips et al., 2004; Vandenabeele et al., 2012) as negative effects were apparent unless tag mass was <1% body mass. Assessing heavy devices is necessarily hampered by their rarity based on common sense and the longstanding adoption of the "5% rule", meaning that <5% of all studies located in our literature search, and <1% retained for analysis, fitted devices even marginally above this threshold. However, given that proportional tag weight had no effect when fitted as a continuous variable, as well as the <1% threshold identified, it seems likely that relationships between tag mass and outcome will be nonlinear. While there are sound ethical and scientific grounds for ensuring that the tags used are as small as possible (Casper, 2009), we do not advocate 1% as the "new standard", although it is clearly a desirable target. Rather, the use of biologgers should always be recognized as a trade-off between the importance of the knowledge gained and the potential deleterious effects caused.

The negative effects of tagging on both survival and reproduction were more apparent in species with flapping flight. A broad categorization into "main flying style" (sensu Watanabe & Nathan, 2016) revealed that flapping species were significantly affected by device attachment, a result not seen in soaring or flightless species. While the latter incorporated fewer studies, the lower energetic costs of soaring (between 2.5 and 9.5 times lower, Pennycuick, 2008) may reduce the impact of tagging, and flightless species, i.e. penguins, likely benefit from the intensive work conducted on ensuring the streamlining of device profiles and optimal positioning (Wilson & McMahon, 2006). Tagging was also negatively associated with survival (95% CRI) and reproduction (80% CRI) of species with longer migration distances. We did not find an effect of deployment duration, which in part may be due to a lack of clarity in reporting (Box 1). Nevertheless, these results suggest that particular caution and consideration be given to the choice of biologging device and attachment type in such species, and again highlights the tradeoff inherent in biologging studies.

Reduced survival of tagged birds was also related to specific attachment methods and positions of the device. Both harness and tailmount attachments were associated with negative effects on survival, and these results are also reflected in the negative positional effects of devices on the back and tail, the natural position for these attachment types. There was also a tendency towards reduced survival, using ponchos and leg bands at a more conservative 80% CRI. Tail-mounted and poncho-based effects were based on small sample sizes and results should be treated with caution. However, harness and leg-mounted designs are the most commonly used attachment methods (Tables S1 and S2), with robust effect size estimates. Indeed, the negative mean effect size associated with harness attachments suggests that such a design may not always be appropriate, particularly for flapping species, despite their current widespread use. In contrast, declines in reproductive success were associated with collar attachments (and thus necessarily neck position). Negative impacts of neck collars may be expected given that basic identification collars can reduce survival (Weegman et al., 2016), and demonstrate the need for caution when using such attachments.

Importantly, our analyses highlight the potential for a cumulative effect of fitting tags. Multivariate meta-analysis revealed a positive correlation between mean effect sizes relating to survival and reproduction—such that if tagging produced deleterious effects on survival,

BOX 1 Key information requirements for assessing device impacts

Questions remain about how to accurately assess the impact of the increasing range of biologgers deployed on wild animals. Meta-analysis provides one means of synthesizing results across studies but also highlights certain shortcomings:

1. It is critical to determine whether the return rates of tagged birds approximate those of untagged ones (Thaxter et al., 2016), yet this is infrequently reported.
2. Relevant data can be only partially reported e.g. mean values without *SD*/*SE*, or results of AIC model selection with no indication of actual effect size.
3. Deployment times are frequently difficult to determine, particularly when devices are attached for longer periods (months-years), and can be conflated with device failure time. This makes it difficult to ascertain any chronic impacts of tagging.
4. Many biologging studies are observational rather than experimental, and may suffer from biases, particularly if devices need to be retrieved to collect data (Weiser et al., 2016). In addition, Authier et al. (2013) suggest that control individuals may not be comparable with tagged individuals as both may represent a non-random sample from the population and present a counterfactual study design to address this.

Consequently, we recommend all biologging studies should, at a minimum, provide clear information on:

1. Study Species
2. Number of devices deployed and individuals tagged (including all instances where the tags failed or individuals did not return)
3. mean mass of study individuals
4. method of attachment used in repeatable detail
5. mass of device(s) deployed
6. total length of tag deployment (particularly if different to the length required to address the specific questions analysed).

it was also likely to hinder reproduction (Figure 3). Consequently, the impact of tagging may be greater than apparent in most studies, which typically focus on single responses.

The absence of any strong phylogenetic signal does suggest that, when tagging a species for the first time, using techniques previously employed on similar species is a practical approach (Sergio et al., 2015). However, evidence of some species-level variation in survival-based effect sizes after accounting for common ancestry signifies that caution is still required. In addition, given the low marginal R^2 values reported for each of meta-regression models it is clear that other important moderator variables remain unidentified. These may relate to both biotic and abiotic effects including climatic influences, differences between specific study environments, and ecological differences at both inter- and intra-specific levels. For example, one could imagine that in harsh environmental conditions the effects of tagging are greater than in benign conditions. In addition, in many cases, the exact age and previous life-history of tagged and control birds were unknown. Consequently, we cannot exclude the role of such unmeasured variables in contributing to heterogeneity in tagging effects between studies, but are necessarily constrained by the variables reported in the original studies.

We found few associations between moderator variables and effect sizes relating to body mass, parental care or foraging duration. Although Barron et al. (2010) reported that tagging led to reductions in body condition, we found little evidence to support this, a result also seen in a recent meta-analysis of tag effects (Costantini & Møller, 2013). Small but significant declines in effect size with time are found in many research fields as a result of initial publication bias against non-significant findings (Jennions & Møller, 2001), although we did not detect such a time trend in the current analysis (see Appendix). Both parental care and foraging trip duration mean effect sizes differed from 0 in our standard meta-analysis, suggesting it is worthwhile monitoring both to assess tagging impacts. In contrast, the lack of an effect on body mass suggests that this may not be a particularly suitable measure for assessing the impacts of tagging.

While we have identified that, on average, there is a negative association between tagging and different life-history traits, given the numerous examples where the benefits gained from biologging are substantial, it is reassuring that effect sizes are generally small, allowing researchers to consider whether the likely negative impacts to individuals may potentially be offset by these gains. For example, working with declining or rare species may bring such ethical trade-offs into particular consideration. However, biologging has brought huge advances in conservation biology, including discovering unknown breeding locations (Rayner et al., 2015) and determining interactions with anthropogenic influences (Bodey et al., 2014), all of which have the potential to enhance environmental protection. Biologging has also improved our understanding of disease transmission (Bengtsson et al., 2016), nutrient transfer (Hussey et al., 2015) and the physiological capabilities of animals (Bishop et al., 2015).

Although we found limited effects of many moderator variables, our ability to detect effects was often hindered by the small sample sizes (and hence low power) involved when separating studies into categories. This also prevented us from including potentially interesting interactions between variables. A key finding is, therefore, that the proportion of studies that provide the complete information necessary to assess the impact of the devices used, either within the published manuscript or associated appendices, remains small (just over one-third). These omissions are easily remedied and we call for future studies to include essential information (Box 1).

Assessing the impacts of fitting biologgers is a critical part of any tagging study, and omission of relevant data hinders assessment of whether identified survival rates (or other key life-history traits) are biased by the attachment of biologging or other marking devices (Authier, Péron, Mante, Vidal, & Grémillet, 2013; Barron et al., 2010; Costantini & Møller, 2013; Sergio et al., 2015). In addition, the importance of variables such as attachment type and position suggest that efforts to alleviate any negative effects of tagging need not focus solely upon fitting smaller, lighter tags. Although we have only analysed the impacts from studies on birds our results have ramifications for biologging studies on many other organisms. In this fast-moving field driven by both technology and scientific enquiry, it is crucial we have as comprehensive an understanding of the impacts of devices as possible. As such, we reissue the recommendation from previous reviews (Barron et al., 2010; Bowlin et al., 2010) that all biologging studies should provide key information in order to facilitate meta-analyses such as those conducted here. This will ensure comprehensive comparisons can be undertaken, facilitating the dissemination of best practice and robust conclusions and generalizations, where applicable. The reporting of this key information (Box 1) will ensure that protocols can be improved rather than simply being transferred between species. Such recommendations can be updated as additional information is acquired from the exponentially increasing number of studies and species as device miniaturization continues.

ACKNOWLEDGEMENTS

T.W.B., I.R.C. and S.B. are supported by ERC Consolidator's grant STATEMIG (310820) awarded to S.B. We thank three anonymous reviewers and the Associate Editor for comments that improved the manuscript. We thank Ted Bedlington for assistance during the writing of the manuscript.

AUTHORS' CONTRIBUTIONS

T.W.B., I.R.C., S.C.V. and S.B. conceived the study. T.W.B., I.R.C., F.B., N.P. and A.S. collated and compiled the dataset. T.W.B. and I.R.C. conducted analyses. T.W.B., I.R.C and S.B. wrote the first draft, with all authors contributing to interpretation of results, editing and approval of the final manuscript.

DATA ACCESSIBILITY

Meta-analysis data and information on phylogenies used available at the Dryad Digital Repository https://doi.org/10.5061/dryad.0rp52 (Bodey et al., 2017).

ORCID

Thomas W. Bodey http://orcid.org/0000-0002-5334-9615

Ian R. Cleasby http://orcid.org/0000-0002-4443-0008

REFERENCES

Adachi, T., Costa, D. P., Robinson, P. W., Peterson, S. H., Yamamichi, M., Naito, Y., & Takahashi, A. (2016). Searching for prey in a three-dimensional environment: Hierarchical movements enhance foraging success in northern elephant seals. *Functional Ecology*, *31*, 361–369.

Anonymous. (2012). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour*, *83*, 301–309.

Authier, M., Péron, C., Mante, A., Vidal, P., & Grémillet, D. (2013). Designing observational biologging studies to assess to causal effect of instrumentation. *Methods in Ecology and Evolution*, *4*, 802–810. https://doi.org/10.1111/2041-210X.12075

Barron, D. G., Brawn, J. D., & Weatherhead, P. J. (2010). Meta-analysis of transmitter effects on avian behaviour and ecology. *Methods Ecology and Evolution*, *1*, 180–187. https://doi.org/10.1111/j.2041-210X.2010.00013.x

Bengtsson, D., Safi, K., Avril, A., Fiedler, W., Wikelski, M., Gunnarsson, G., & Waldenström, J. (2016). Does influenza a virus infection affect movement behaviour during stopover in its wild reservoir host? *Royal Society Open Science*, *3*, 150633. https://doi.org/10.1098/rsos.150633

BirdLife International. (2016). IUCN Red list for birds. Downloaded from http://www.birdlife.org on 27/11/2016.

Bishop, C. M., Spivey, R. J., Hawkes, L. A., Batbayar, N., Chua, B., Frappell, P. B., … Butler, P. J. (2015). The roller coaster flight strategy of bar-headed geese conserves energy during Himalayan migrations. *Science*, *347*, 250–254. https://doi.org/10.1126/science.1258732

Bodey, T. W., Cleasby, I. R., Bell, F., Parr, N., Schultz, A., Votier, S. C., & Bearhop, S. (2017). Data from: A phylogenetically controlled meta-analysis of biologging device effects on birds: Deleterious effects and a call for more standardized reporting of study data. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.0rp52

Bodey, T. W., Jessopp, M. J., Votier, S. C., Gerritsen, H. D., Cleasby, I. R., Hamer, K. C., … Bearhop, S. (2014). Seabird movement reveals the ecological footprint of fishing vessels. *Current Biology*, *24*, R514–R515. https://doi.org/10.1016/j.cub.2014.04.041

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons. https://doi.org/10.1002/9780470743386

Bowlin, M. S., Henningsson, P., Muijres, F. T., Vleugels, R. H. E., Liechti, F., & Hedenström, A. (2010). The effects of geolocator drag and weight on the flight ranges of small migrants. *Methods in Ecology and Evolution*, *1*, 398–402. https://doi.org/10.1111/j.2041-210X.2010.00043.x

Casper, R. M. (2009). Guidelines for the instrumentation of wild birds and mammals. *Animal Behaviour*, *78*, 1477–1483. https://doi.org/10.1016/j.anbehav.2009.09.023

Cleasby, I. R., & Nakagawa, S. (2012). The influence of male age on within-pair and extra-pair paternity in passerines. *Ibis*, *154*, 318–324. https://doi.org/10.1111/j.1474-919X.2011.01209.x

Costantini, D., & Møller, A. P. (2013). A meta-analysis of the effects of geolocator application on birds. *Current Zoology*, *59*, 697–706. https://doi.org/10.1093/czoolo/59.6.697

Dunning, J. B. Jr. (2008). *CRC Handbook of avian body masses* (2nd ed.). Boca Raton, FL: CRC Press.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Gannon, W. L., & Sikes, R. S. (2007). Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *Journal of Mammalogy*, *88*, 809–823.

Godfrey, J. D., & Bryant, D. M. (2003). Effects of radio transmitters: Review of recent radio-tracking studies. *Science for Conservation*, *214*, 83–95.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*, 1–22.

Hadfield, J. D., & Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, *23*, 494–508. https://doi.org/10.1111/j.1420-9101.2009.01915.x

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., … Whoriskey, F. G. (2015). Aquatic animal telemetry: A panoramic window into the underwater world. *Science*, *348*, 1255642. https://doi.org/10.1126/science.1255642

Jennions, M. D., Lortie, C. J., Rosenberg, M. S., & Rothstein, H. R. (2013). Publication and related biases. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution* (pp. 207–236). Princeton, NJ: University Press, Princeton.

Jennions, M. D., & Møller, A. P. (2001). Relationships fade with time: A meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society: Series B*, *269*, 43–48.

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*, 444–448. https://doi.org/10.1038/nature11631

Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*, 1222–1232.

Kenward, R. (2001). *A manual for wildlife radio tracking* (2nd Ed.). Cambridge, MA: Academic Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Beverly Hills, CA: Sage Publishing.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337. https://doi.org/10.1023/A:1008929526011

Mauck, R. A., & Grubb, T. C., Jr (1995). Petrel parents shunt all experimentally increased reproductive costs to their offspring. *Animal Behaviour*, *49*, 999–1008.

Milner-Gulland, E. J., Fryxell, J. M., & Sinclair, A. R. E. (2011). *Animal migration – A synthesis*. Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199568994.001.0001

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Review*, *82*, 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolution Ecology*, *26*, 1253–1274. https://doi.org/10.1007/s10682-012-9555-5

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods of Ecological Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Pennycuick, C. J. (2008). *Modelling the flying bird*. London: Academic Press.

Phillips, R. A., Silk, J. R. D., Croxall, J. P., Afanasyev, V., & Briggs, D. R. (2004). Accuracy of geolocation estimates for flying seabirds. *Marine Ecology Progress Series*, *266*, 265–272. https://doi.org/10.3354/meps266265

R Development Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rayner, M. J., Gaskin, C. P., Fitzgerald, N. B., Baird, K. A., Berg, M. M., Boyle, D., … Ismar, S. M. H. (2015). Using miniaturised radio telemetry to discover the breeding grounds of the endangered New Zealand Storm Petrel *Fregetta maoriana*. *Ibis*, *157*, 754–766. https://doi.org/10.1111/ibi.12287

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S., & Nakagawa, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: Its magnitude and implications. *Ecology*, *97*, 3293–3299.

Sergio, F., Tavecchia, G., Tanferna, A., López Jiménez, L., Blas, J., De Stephanis, R., … Hiraldo, F. (2015). No effect of satellite tagging on survival, recruitment, longevity, productivity and social dominance of a raptor, and the provisioning and condition of its offspring. *Journal of Applied Ecology*, *52*, 1665–1675. https://doi.org/10.1111/1365-2664.12520

Thaxter, C. B., Ross-Smith, V. H., Clark, J. A., Clark, N. A., Conway, G. J., Masden, E. A., … Burton, N. H. K. (2016). Contrasting effects of GPS device and harness attachment on adult survival of Lesser Black-backed Gulls *Larus fuscus* and Great Skuas *Stercorarius skua*. *Ibis*, *158*, 279–290. https://doi.org/10.1111/ibi.12340

Vandenabeele, S. P., Shepard, E. L., Grogan, A., & Wilson, R. P. (2012). When three per cent may not be three per cent; device-equipped seabirds experience variable flight constraints. *Marine Biology*, *159*, 1–14. https://doi.org/10.1007/s00227-011-1784-6

Wakefield, E. D., Bodey, T. W., Bearhop, S., Blackburn, J., Colhoun, K., Davies, R., … Hamer, K. C. (2013). Space partitioning without territoriality in gannets. *Science*, *341*, 68–70. https://doi.org/10.1126/science.1236077

Watanabe, Y. Y., Goldman, K. J., Caselle, J. E., Chapman, D. D., & Papastamatiou, Y. P. (2015). Comparative analyses of animal-tracking data reveal ecological significance of endothermy in fishes. *Proceedings of the National Academy of Science of the United States of America*, *112*, 6104–6109. https://doi.org/10.1073/pnas.1500316112

Watanabe, Y. Y., & Nathan, R. (2016). Flight mode affects allometry of migration range in birds. *Ecology Letters*, *19*, 907–914. https://doi.org/10.1111/ele.12627

Weegman, M. D., Bearhop, S., Fox, A. D., Hilton, G. M., Walsh, A. J., Mcdonald, J. L., & Hodgson, D. J. (2016). Integrated population modelling reveals a perceived source to be a cryptic sink. *Journal of Animal Ecology*, *85*, 467–475. https://doi.org/10.1111/1365-2656.12481

Weiser, E. L., Lanctot, R. B., Brown, S. C., Alves, J. A., Battley, P. F., Bentzen, R., … Sandercock, B. K. (2016). Effects of geolocators on hatching success, return rates, breeding movements, and change in body mass in 16 species of Arctic-breeding shorebirds. *Movement Ecology*, *4*, 12. https://doi.org/10.1186/s40462-016-0077-6

Wilson, R. P., & McMahon, C. R. (2006). Measuring devices on wild animals: What constitutes acceptable practice? *Frontiers of Ecology & Environment*, *4*, 147–154. https://doi.org/10.1890/1540-9295(2006)004[0147:MDOWAW]2.0.CO;2

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Bodey TW, Cleasby IR, Bell F, et al. A phylogenetically controlled meta-analysis of biologging device effects on birds: Deleterious effects and a call for more standardized reporting of study data. *Methods Ecol Evol*. 2018;9:946–955. https://doi.org/10.1111/2041-210X.12934