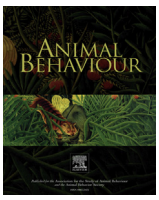
Animal Behaviour 222 (2025) 123102

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav



Commentary

Repeatability and intraclass correlations from time-to-event data: towards a standardized approach

Kelsey B. McCune [a, *], Coralie Williams [b, c], Ned A. Dochtermann [d], Holger Schielzeth [e], Shinichi Nakagawa [b, f, *]

[a] College of Forestry, Wildlife and Environment, Auburn University, AL, U.S.A.
[b] Evolution & Ecology Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia
[c] School of Mathematics and Statistics, University of New South Wales, Sydney, Australia
[d] Department of Biological Sciences, North Dakota State University, Fargo, ND, U.S.A.
[e] Institute for Ecology and Evolution, Friedrich Schiller University Jena, Thuringia, Germany
[f] Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

ARTICLE INFO

Article history:
Received 20 July 2024
Initial acceptance 17 October 2024
Final acceptance 27 November 2024
Available online xxx
MS. number: AS-24-00456

Keywords:
consistent individual differences
latency
repeatability
survival analysis

Many biological features are expressed as 'time-to-event' traits, such as time to first reproduction or time to first response to some stimulus. The analysis of these traits frequently produces right-censored data in cases where no event has occurred within a certain time frame. The Cox proportional hazards (CPH) model, a type of survival analysis, accounts for censored data by estimating the hazard of an event occurring at each time point. While random effect variances can be estimated in CPH models, it is currently not possible to estimate within-cluster variance. Consequently, we lack a general method for calculating ecologically and evolutionary relevant variances and metrics like repeatability from time-to-event data. We here present a solution to this issue. We first describe the characteristics of CPH models and introduce repeatability as an intraclass correlation coefficient (ICC). We demonstrate how CPH models with discrete time intervals are comparable to binomial generalized linear mixed-effects models (GLMMs) with the complementary log-log link. Through this equivalence, we show how to estimate an ICC using the estimates of the random effects variance component(s) resulting from CPH models and the distribution-specific variance (within-cluster variance) from the binomial GLMM. We provide a case study and online materials to demonstrate how our new method for ICC for time-to-event data can be implemented and used. We conclude that the proposed method will not only generate a standard way to quantify consistent individual differences (ICC) from time-to-event data, but also broaden the use of survival analysis outside of the typical implementation for survivorship studies.

Consistent individual differences, where traits are relatively fixed within individuals either genetically or during development, can have significant ecological and evolutionary consequences (Dochtermann & Dingemanse, 2013; Réale et al., 2007; Wolf & Weissing, 2012). For example, individual differences can impact dispersal (Cote et al., 2010), range expansion (Duckworth & Badyaev, 2007) and persistence in the face of rapid environmental change (Lapiedra et al., 2017; Wright et al., 2010). The identification of consistency (i.e. repeatability) in a trait requires repeated measures within and across individuals to quantify the proportion of the total (sample) variance that is attributable to differences among individuals relative to variation in the trait within individuals (Dingemanse & Dochtermann, 2013; Nakagawa & Schielzeth, 2010). Consequently, researchers have developed rigorous methods and statistical techniques to meet the growing interest in quantifying consistent individual differences in behaviour (Dingemanse & Wright, 2020; Stoffel et al., 2017).

Despite significant progress, methods to quantify individual differences often result in data that are difficult to analyse, or inappropriately used, with the current statistical tools. Time-to-event data, such as the latency to respond to a stimulus, approach a threat or solve a problem, are widespread across studies of animal behaviour and cognition, as well as broadly in ecology and evolution (Table 1). These data have unique features in that they are time dependent and often include right-censored values (censoring) when an event does not occur within the experimental or observational time frame (Machin et al., 2006). Researchers often

* Corresponding authors.
E-mail addresses: kelseybmccune@gmail.com (K. B. McCune), s.nakagawa@unsw.edu.au, snakagaw@ualberta.ca (S. Nakagawa).

https://doi.org/10.1016/j.anbehav.2025.123102

assign ceiling values (e.g. the maximum duration of the trial) to trials where individuals never produced the event and then analyse the data using random effects models that assume a Poisson or Gaussian distribution (e.g. Johnstone & Garvey, 2023; Lukas et al., 2021; Peignier et al., 2022; Vámos & Shaw, 2024). However, it is problematic to assign such arbitrary response values to the individuals where the event was not observed. Failing to account for the right-censored nature of the data could bias results as the upper end of the range of performance is truncated. Statistical solutions are needed, because it is often logistically not feasible to give all individuals unlimited trial durations.

An alternative tool for the analysis of time-to-event data is survival analysis. Survival analysis, such as Cox proportional hazards regression (Cox, 1972), accounts for time-dependent and right-censored data. Although primarily used in biomedical research, use in behaviour and ecology is increasing (e.g. Barak et al., 2018; Griffin & Diquelou, 2015; McCune et al., 2022; van den Heuvel et al., 2023; Table 1). However, there are currently no widely known methods to quantify repeatability as the intraclass correlation coefficient (ICC), which uses variance components from survival analyses.

In this paper, we describe how to quantify the ICC from time-to-event data. First, we introduce the statistical features of the Cox proportional hazards (CPH) model and define the ICC. Next, we show that CPH models with discrete time intervals are analogous to binomial generalized linear mixed-effects models (GLMM) by describing a method for restructuring time-to-event data for use in the GLMM framework. This equivalence provides an accessible pathway to estimate ICC. Finally, we provide a worked example using our new method on real-world data. We present several additional worked examples with real data as well as a small simulation study in an online tutorial at https://kelseybmccune.github.io/Time-to-Event_Repeatability/Online-tutorial.html.

PROPORTIONAL HAZARDS MODELS AND THE INTRACLASS CORRELATION COEFFICIENT

Cox proportional hazards models (or Cox regression) estimate the hazard of an event occurring in relation to predictor variables with time-to-event data (i.e. the time taken to an event or censoring; Cox, 1972). The hazard is a rate (or risk) of an event occurring at time t. The hazard rate is defined in a Cox model as:

$$\lambda_i(t) = \lambda_0(t)\exp(b_1x_1 + b_2x_2 + \ldots + b_mx_m), \quad (1)$$

where $\lambda(t)$ is the hazard rate at time t for the ith subject (or observation), $\lambda_0(t)$ is the baseline hazard rate and $\beta_1, \beta_2, \ldots, \beta_m$ are the regression coefficients for the predictor variables $x_1, x_2, \ldots, x_m$. Notably, $\lambda_0(t)$ takes the place of an intercept as $\exp(\ln(\lambda_0(t)) + b_1x_1 + \ldots)$ or $\exp(b_0 + b_1x_1 + \ldots)$, where $\ln(\lambda_0(t)) = b_0$ ('ln' is a natural logarithm). Equation (1) can be rearranged to:

$$\ln\left(\frac{\lambda_i(t)}{\lambda_0(t)}\right) = b_1x_1 + b_2x_2 + \ldots + b_mx_m, \quad (2)$$

for the right-hand side to take a linear form, which is more familiar for many readers due to the similarity to linear regression, although it has neither an intercept (i.e. b_0) nor a residual term (i.e. ε_i). To fit such a model using, for example, the R statistical language (R Core Team, 2023), one needs to provide the time-to-event data in the form of a 'Surv' object (e.g. 'Surv (time, event)', where 'time' is the time taken to an event and 'event' indicates whether the event was observed or censored, coded as 0 or 1). The Cox model can be fitted using the 'coxph' function in the 'survival' package (Therneau, 2024).

Now let us assume that we have a single predictor variable, sex (x_{sex}), for a time-to-event data set (e.g. to study sex-specific latency to solve a task), and we have a single random effect (or cluster) α (e.g. individuals or populations). The Cox proportional hazards model can be extended to include a random effect (like individual identity), which is often referred to as the 'frailty' term and a Cox regression with a single random effect is therefore known as the frailty model:

$$\ln\left(\frac{\lambda_{ij}(t)}{\lambda_0(t)}\right) = b_{sex}x_{sex} + \alpha_i, \quad (3)$$

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right),$$

where $\lambda_{ij}(t)$ is the hazard rate at time t for the ith individual for the jth occasion (observation). Such a frailty model can be fitted using the 'coxme' function in the R package 'coxme' (Therneau, 2022) as well as with the 'coxph' function.

Now that we have defined the Cox model, let us define the repeatably or intraclass correlation coefficient (ICC) in its simplest form when the trait of interest (the response variable) is a Gaussian variable (i.e. a model that has normally distributed residuals):

$$\text{ICC} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}, \quad (4)$$

where σ_α^2 is the variance of the random effect (the between-cluster variance, where a cluster could be individual identity) and σ_ϵ^2 is the variance of the residuals (or within-cluster variance; Nakagawa & Schielzeth, 2010). The ICC can be interpreted as the proportion of the total variance that is explained by between-cluster variance.

Table 1
Examples of the use of time-to-event data with cluster variable(s) from behaviour and ecology

Topic	Time-to-event variables tested	Cluster variable	Example studies
Animal personality	Latency to approach simulated intruder or predator	Individual ID	Holzmann and Cordóba (2024); Peignier et al. (2022)
	Latency to forage	Individual ID, Year ID	Vámos and Shaw (2024)
	Latency to emerge from a shelter	Individual ID	Lapiedra et al. (2017); Mazué et al. (2015)
Animal cognition	Number of trials to reach criterion	Individual ID	McCune et al. (2023); van den Heuvel et al. (2023)
	Latency to solve a task	Individual ID, Treatment ID	Griffin and Diquelou (2015); Lenkei et al. (2020)
Animal ecology	Seed dispersal distance (animal mediated)	Individual ID, Site ID	Brehm et al. (2019); Wrobel et al. (2022)
	Distance to detect camouflaged animal	Individual ID, Site ID	Briolat et al. (2021)
	Distance to dispersal	Individual ID	Villegas-Ríos et al. (2017)
	Latency to initiate nesting	Site ID, Social group ID	Brandl et al. (2019); Stroeymeyt et al. (2011)
	Latency to migrate	Individual ID	Ramírez et al. (2016)
Plant phenology	Survival	Treatment ID	Waser et al. (2000)
	Latency to flowering	Site ID	de Manincor et al. (2023); Villagomez et al. (2021)
	Latency to initiate pollinator activity	Site ID	de Manincor et al. (2023); Villagomez et al. (2021)

In many situations, including time-to-event data, the Gaussian distribution is not appropriate, and the residuals are not normally distributed. Methods to estimate within-cluster variance have previously been developed for some non-Gaussian data that can inform the calculation of ICC from a CPH model (Nakagawa & Schielzeth, 2010). The ICC can be calculated for generalized linear mixed-effect models (GLMMs) with non-Gaussian error distributions and link functions other than identity links. For example, the R package 'rptR' can be used to calculate ICC from various GLMMs (Stoffel et al., 2017), via the 'lmer' and 'glmer' functions in the 'lme4' package (Bates et al., 2015). Nakagawa and Schielzeth (2010) suggested that for non-Gaussian models, the within-cluster variance is determined by the distributional assumptions of a GLMM. For binomial models with the logit link, for example, σ^2_ε can be assumed to be $\pi^2/3$ (where σ^2_ε is called the distributional specific variance; Nakagawa & Schielzeth, 2010); $\pi^2/3$ is the variance of the assumed underlying distribution, i.e. the logistic distribution. Published formulae are available for all common GLMMs and can generally be derived by the delta method for other GLMM families (Nakagawa et al., 2017).

Cox models do not make any distributional assumptions about the hazard rate; it is an inherently nonparametric analysis (equations 1–2); more precisely, these models do not make any assumptions about residuals or within-subject variability. Indeed, although frailty models (equation 3) have a random effect term with a Gaussian or gamma distribution, they do not make distributional assumptions about the residual deviations, so that the frailty model is referred to as semiparametric. Given this lack of distributional assumptions, we cannot calculate the ICC for Cox models using these current tools, because the residual variance is missing. However, a formula for the nonparametric (or likely more accurately, semiparametric) version of ICC (ICC_{np}) for the frailty model is known when the random effect is assumed to be gamma-distributed on the exponential scale. If we denote the variance from a gamma distribution as θ_α under equation (3), then ICC_{np} can be written as:

$$ICC_{np} = \frac{\theta_\alpha}{\theta_\alpha + 2}, \tag{5}$$

$$\exp(\alpha_i) \sim G\left(\frac{1}{\theta_\alpha}, \frac{1}{\theta_\alpha}\right),$$

where the first $1/\theta_\alpha$ and the second $1/\theta_\alpha$, are the shape and the rate parameter of the gamma distribution, respectively (such parameterization results in the mean, $E(\exp(\alpha_i)) = 1$, and variance, $\mathrm{Var}(\exp(\alpha_i)) = \theta_\alpha$).

The estimate ICC_{np} represents Kendall's τ, that is, the rank correlation or concordance for within-cluster observations for the frailty model (Hougaard, 2000). Unfortunately, there is no closed-form formula when assuming a Gaussian distribution for the random effect, as in equation (3). Nevertheless, ICC_{np} can be obtained numerically, and we provide an R function based on the tau function from the R package 'parfm' (Munda et al., 2012) in section 2.1.2 of the online tutorial (see ICC from coxme with an exploded data set). We note that σ^2_α (Gaussian) and θ_α (gamma) are unlikely to be the same, but the two ICC_{np} values under two different assumptions (Gaussian and gamma) are often very similar (see simulations in section 4 of the online tutorial).

An issue with the ICC_{np} is that it is not a parametric version of ICC, and its comparison to typical parametric ICCs is not straightforward. More importantly, it is not clear whether this method can be extended to a Cox model that has more than one random effect (at least, practically speaking). However, we can solve both issues by restructuring the time-to-event data of the Cox model into a data set where we can fit a GLMM and then obtain a parametric version of ICC via distributional-specific (residual) variance formulae from GLMMs.

COX PROPORTIONAL HAZARDS MODELS AND GENERALIZED LINEAR MIXED-EFFECTS MODELS

In the statistical literature, it has been shown that the frailty model (equation 3) can be fitted as a Poisson GLMM (known as the piecewise exponential model; e.g. Hirsch et al., 2016) or a binomial GLMM (the discrete time model: Finkelstein, 1986; Suresh et al., 2022; for an accessible overview, see Austin, 2017). Here, we show how a discrete time model, more specifically a binomial GLMM with the complementary log-log (cloglog) link can be used to fit a comparable model as equation (3) by 'exploding' the time-to-event data into arbitrary discrete time intervals (Fig. 1 shows an example of such an exploded data set compared to the original). Briefly, within the range of a continuous time-to-event measure (e.g. latency to approach), discrete time intervals are chosen of a specific duration. Within each interval, it is then evaluated whether an individual (or other cluster variables) experienced the event (event = 1) or not (event = 0). In section 2.1.2 of the online tutorial, we demonstrate how to create an exploded data set from the classic time-to-event data. It is not necessary to consider biological factors in choosing the number of intervals (and, therefore, the duration of each interval). In section 4 of the online tutorial, we show using simulations that the number and duration of time intervals do not impact the estimates of the fixed and random effect variances. If we assume we have three (arbitrary discrete) time intervals ($t1$, $t2$ and $t3$), this binomial GLMM (without an intercept) is defined as:

$$\ln\left(\frac{-\ln\left(1-\lambda_{ijk}(t)\right)}{-\ln(1-\lambda_{0k}(t))}\right) = b_{t1}x_{t1} + b_{t2}x_{t2} + b_{t3}x_{t3} + b_{\mathrm{sex}}x_{\mathrm{sex}} + \alpha_i \tag{6}$$

where $\lambda_{ijk}(t)$ is the hazard rate at the time t for the ith subject at the jth occasion in the kth time interval ($k = t1, t2, t3$), $\lambda_{0k}(t)$ is the baseline hazard rate for the kth time interval, x_{t1}, x_{t2}, x_{t3} are the indicator variables for the time intervals, and b_{t1}, b_{t2}, b_{t3} are the regression coefficients for the time intervals (the population-average hazard rates at times $t1$, $t2$ and $t3$). Note that the cloglog link is $\ln(-\ln(1-\lambda))$, where λ is the rate at which the event occurs. Thus, the left-hand side of equation (6) consists of the cloglog-transformed hazard rate ($\lambda_{ijk}(t)$) and baseline hazard rate ($\lambda_{0k}(t)$).

Rather remarkably, b_{sex} and σ^2_α in equation (6) are estimated to be the same as those in equation (3) despite the very different data structures for the two models (i.e. time-to-event data versus exploded data; Fig. 1). Note that we show in the online tutorial (section 4) the equivalence of b_{sex} and σ^2_α between the Cox (frailty) model, fitted with 'coxph' and 'coxme', and the binomial GLMM, fitted with 'glmer' with event (0 or 1) as the response.

Given the equivalence of the regression parameters, equation (6) will not typically need to be fit. We can use variance components obtained from Cox models to estimate ICC under the assumptions of a binomial GLMM with the complementary log-log link. Under this GLMM, the distribution-specific variance σ^2_ε (as used in equation 4) is $\pi^2/6$ on the latent scale. This means we can define ICC for equations (3) and (6) as (Nakagawa et al., 2017):

Original data set

ID	Trial	Latency
A	1	22
A	2	24
B	1	43
B	2	50*
C	1	8
C	2	12

* Ceiling value

Discrete time interval ('exploded') data set

ID	Trial	Interval	Time start	Time stop	Event
A	1	1	0	10	0
A	1	2	10	20	0
A	1	3	20	22	1
B	1	1	0	10	0
B	1	2	10	20	0
B	1	3	20	30	0
B	1	4	30	40	0
B	1	5	40	43	1
C	1	1	0	8	1
A	2	1	0	10	0
A	2	2	10	20	0
A	2	3	20	24	1
B	2	1	0	10	0
B	2	2	10	20	0
B	2	3	20	30	0
B	2	4	30	40	0
B	2	5	40	50*	0
C	2	1	0	10	0
C	2	2	10	12	1

Figure 1. Demonstration of a simple time-to-event clustered data set in the traditional form (left) and the 'exploded' form (right). Time intervals ('Time start' to 'Time stop') replace the 'Latency' variable, and an 'Event' column is created that notes whether that individual achieved the event in that interval and trial. For instance, trial 1 for individual B is exploded into five time intervals with cut points every 10 s, until B achieved the event at 43 s.

$$\mathrm{ICC} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \pi^2/6}. \tag{7}$$

In Fig. 2, we show the parametric version of ICC and the nonparametric version ($\mathrm{ICC_{np}}$) are well correlated but not equivalent (analogous to the relationship between Pearson's r and Kendall's τ). We prefer the use of ICC as in equation (7) over $\mathrm{ICC_{np}}$ because the parametric version is more comparable to other ICC estimates derived from GLMMs that are commonly used in ecology and evolution (Nakagawa et al., 2017).

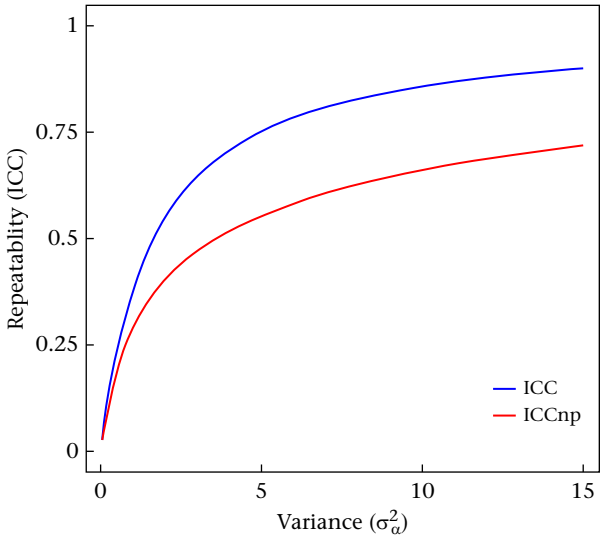


Figure 2. Repeatability estimates from the parametric equation for intraclass correlation coefficient (ICC) (in blue) compared to the nonparametric ICC (in red) obtained from Kendall's τ. Repeatability values from the two estimates are correlated, but not identical.

Furthermore, the advantage of this approach is that we can add more than one random effect. For example, imagine we have additional levels of clustering (such as population identity) modelled as random effects. Adding those additional random effects to equation (3) yields:

$$\ln\left(\frac{\lambda_{ijl}(t)}{\lambda_0(t)}\right) = b_{\mathrm{sex}}x_{\mathrm{sex}} + \alpha_i + \gamma_l, \tag{8}$$

$$\gamma_l \sim N\left(0, \sigma_\gamma^2\right),$$

where γ_l is the random effect for the l th level of the second cluster, which is assumed to be normally distributed with the mean of zero and the variance of σ_γ^2 (although Cox models with more than one random factor cannot be fitted with the 'coxph' function but can be fitted with the 'coxme' function). It is important to notice that the two random effects can be 'nested' or 'crossed' and the difference affects interpretation but not the model fitting (provided the data are coded appropriately; Schielzeth & Nakagawa, 2013).

An example of the nested random effects is individual (α_i) and population (γ_l) where individuals are nested within populations. In this case, the ICC for individuals can be defined as:

$$\mathrm{ICC_{ind1}} = \frac{\sigma_\alpha^2 + \sigma_\gamma^2}{\sigma_\alpha^2 + \sigma_\gamma^2 + \pi^2/6}. \tag{9}$$

The reason both variance components are included in the numerator of the ICC for individuals is that some of the individual consistency comes from individuals belonging to specific populations. Yet, one may be purely interested in the individual σ_α^2 without the effect of population σ_γ^2. If so, they need to adjust equation (9) accordingly. Incidentally, if we do not include the population random effects in a GLMM, equation (7) will give the same ICC value as equation (9) with both individual and population random effects.

An example of the crossed random effects is individual (α_i) and year (γ_l) where individuals are not nested within years but observed across multiple years. This time, the ICC for individuals can be written as:

$$\mathrm{ICC}_{\mathrm{ind2}} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\gamma^2 + \pi^2/6}. \quad (10)$$

If one wants to remove or adjust for the effect of year, then ICC for individuals simplifies to equation (7). Indeed, all the ICC formulas presented above represent 'adjusted' repeatabilities (ICC) sensu Nakagawa and Schielzeth (2010) because the effect of a fixed effect (sex, in our example) is accounted for in the models. We can obtain 'unadjusted' repeatability (ICC) by fitting the model without the fixed effect (sex), for example by changing equation (3):

$$\ln\left(\frac{\lambda_{ij}(t)}{\lambda_0(t)}\right) = \alpha_i. \quad (11)$$

Importantly, the model without any fixed effects (e.g. equation 11) should give an ICC equal to or larger than the model with fixed effects (e.g. equation 3). We show how to fit the models and obtain the ICC estimates we described above at https://kelseybmccune.github.io/Time-to-Event_Repeatability/Online-tutorial.html.

CASE STUDY

We used data from a study comparing the performance of Mexican jays, *Aphelocoma wollweberi*, on a multi-access puzzle box in captivity ($N = 10$ individuals) and in the wild ($N = 7$; McCune et al., 2019). The repeated time-to-event measure consisted of the latency to access food from each of the four different access options (loci). It was not the goal of the original study to evaluate the individual repeatability of solving performance, but that is what we focused on here. Around 44% of the data were censored as not all jays solved one or more of the puzzle box loci in the allotted experimental time frame. We used the 'coxme' function (Therneau, 2022) to model the latency to solve each locus on the puzzle box as a function of treatment (wild or captive jay; a fixed effect) with individual identity as a random effect. This analysis results in an adjusted (accounting for the fixed effect of wild/captive) variance estimate (as in equation 3) of 3.16. Using equation (7), we estimate the ICC for individual identity as 0.66 (CI: 0.38–0.87; $P < 0.01$). However, if we are not interested in the effect of the treatment as a part of among-individual variation of solving latency (i.e. equation 11), then we get the unadjusted random effect variance and an ICC estimate of 0.74 (CI: 0.49–0.90; $P < 0.01$) from equation (7). For the data and code, including functions to estimate the *P* values and 95% confidence intervals of ICC, see section 3 of the online tutorial.

CONCLUSION

Accurate and standardized ways of estimating consistent individual differences are important for answering fundamental questions in animal behaviour, evolution and ecology (e.g. invasive species: Carere & Gherardi, 2013; ecosystem services: Zwolak, 2018). The techniques in this paper provide a standardized and potentially more accurate method for repeatability estimates from time-to-event data, one of the most commonly collected measures in animal behaviour (Table 1; also see Réale et al., 2007; Takola et al., 2021). There are a few assumptions that data must meet for the use of Cox proportional hazards models. We note that researchers including fixed effects in their models should check that they do not violate the proportional hazards assumption of a constant hazard ratio across time (Machin et al., 2006). We refer the reader to Therneau et al. (2024) for guidance on using time-dependent covariates in the Cox model if this assumption is violated. We recognize that survival analyses, or methods for accounting for censored data in parametric analyses, can be implemented through a Bayesian framework, for example, using the 'brms' package in R (Bürkner, 2017). However, we focused here on a frequentist analysis through the R functions 'coxph' and 'coxme' to increase the ease of implementation, especially for those already familiar with the 'rptR' package. Moreover, the approach and formula we provide for calculating ICC can also be applied in a Bayesian context and, arguably, more appropriately represent many behavioural responses. Finally, many behavioural and evolutionary ecologists may believe Cox regression and related models are only for 'survival' analyses. However, there are many different uses without collecting data across the entire lifetime of the subject (e.g. time to fledge, where the cluster variable is nest identity). We hope that Cox and related regression analyses will be more widely used to address a more general range of questions in the future.

Author Contributions

Ned A. Dochtermann: Writing – review & editing, Validation, Conceptualization. **Kelsey B. McCune:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Shinichi Nakagawa:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Methodology, Investigation. **Holger Schielzeth:** Writing – review & editing, Validation. **Coralie Williams:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation.

Data Availability

The data and code for this study are available at https://github.com/kelseybmccune/Time-to-Event_Repeatability/.

Declaration of Interest

The authors declare no conflicts of interest.

Acknowledgments

This project was made possible through a Statistical Quantification of Individual Differences (SQuID) travelling research fellowship awarded to K.M. from the Research Council of Norway INTPART project, grant number 309356 to Jonathan Wright. We are grateful for Josep Carrasco's statistical advice and helpful discussions with Martin Stoffel. S.N. was supported by the Australian Research Council Discovery Grant (DP230101248) and the Canada Excellence Research Chair Program (CERC-2022-00074). H.S. was supported by the German Research Foundation (DFG, grant 316099922).

References

Austin, P. C. (2017). A tutorial on multilevel survival analysis: Methods, models and applications. *International Statistical Review, 85*, 185–203. https://doi.org/10.1111/insr.12214

Bürkner, P. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Barak, R. S., Lichtenberger, T. M., Wellman-Houde, A., Kramer, A. T., & Larkin, D. J. (2018). Cracking the case: Seed traits and phylogeny predict time to germination in prairie restoration species. *Ecology and Evolution, 8*, 5551–5562. https://doi.org/10.1002/ece3.4083

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Brandl, H. B., Griffith, S. C., & Schuett, W. (2019). Wild zebra finches choose neighbours for synchronized breeding. *Animal Behaviour, 151*, 21–28.

Brehm, A. M., Mortelliti, A., Maynard, G. A., & Zydlewski, J. (2019). Land-use change and the ecological consequences of personality in small mammals. *Ecology Letters, 22*(9), 1387–1395.

Briolat, E. S., Arenas, L. M., Hughes, A. E., Liggins, E., & Stevens, M. (2021). Generalist camouflage can be more successful than microhabitat specialisation in natural environments. *BMC Ecology and Evolution, 21*, Article 151.

Carere, C., & Gherardi, F. (2013). Animal personalities matter for biological invasions. *Trends in Ecology & Evolution, 28*, 5–6. https://doi.org/10.1016/j.tree.2012.10.006

Cote, J., Clobert, J., Brodin, T., Fogarty, S., & Sih, A. (2010). Personality-dependent dispersal: Characterization, ontogeny and consequences for spatially structured populations. *Philosophical transactions of the Royal Society of London, Series B, Biological Sciences, 365*, 4065–4076. https://doi.org/10.1098/rstb.2010.0176

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society - Series B: Statistical Methodology, 34*, 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

de Manincor, N., Fisogni, A., & Rafferty, N. E. (2023). Warming of experimental plant–pollinator communities advances phenologies, alters traits, reduces interactions and depresses reproduction. *Ecology Letters, 26*(2), 323–334.

Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: Mixed-effect modelling approaches. *Journal of Animal Ecology, 82*, 39–54. https://doi.org/10.1111/1365-2656.12013

Dingemanse, N. J., & Wright, J. (2020). Criteria for acceptable studies of animal personality and behavioural syndromes. *Ethology, 126*, 865–869. https://doi.org/10.1111/eth.13082

Dochtermann, N. A., & Dingemanse, N. J. (2013). Behavioral syndromes as evolutionary constraints. *Behavioral Ecology, 24*, 806–811. https://doi.org/10.1093/beheco/art002

Duckworth, R. A., & Badyaev, A. V. (2007). Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 1–6.

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics, 42*, 845–854. https://doi.org/10.2307/2530698

Griffin, A. S., & Diquelou, M. C. (2015). Innovative problem solving in birds: A cross-species comparison of two highly successful passerines. *Animal Behaviour, 100*, 84–94. https://doi.org/10.1016/j.anbehav.2014.11.012

Hirsch, K., Wienke, A., & Kuss, O. (2016). Log-normal frailty models fitted as Poisson generalized linear mixed models. *Computer Methods and Programs in Biomedicine, 137*, 167–175. https://doi.org/10.1016/j.cmpb.2016.09.009

Holzmann, I., & Córdoba, R. S. (2024). Individual vocal recognition and dear enemy effect in the black-and-gold howler monkey (*Alouatta caraya*). *Behavioral Ecology and Sociobiology, 78*(4), Article 51.

Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer.

Johnstone, K. C., & Garvey, P. M. (2023). To boldly go: Methods to quantify personality in mustelids. *Animal Behaviour, 202*, 139–147. https://doi.org/10.1016/j.anbehav.2023.05.010

Lapiedra, O., Chejanovski, Z., & Kolbe, J. J. (2017). Urbanization and biological invasion shape animal personalities. *Global Change Biology, 23*, 592–603. https://doi.org/10.1111/gcb.13395

Lenkei, R., Faragó, T., Kovács, D., Zsilák, B., & Pongrácz, P. (2020). That dog won't fit: Body size awareness in dogs. *Animal Cognition, 23*(2), 337–350.

Lukas, J., Kalinkat, G., Miesen, F. W., Landgraf, T., Krause, J., & Bierbach, D. (2021). Consistent behavioral syndrome across seasons in an invasive freshwater fish. *Frontiers in Ecology and Evolution, 8*, Article 583670. https://doi.org/10.3389/fevo.2020.583670

Machin, D., Cheung, Y. B., & Parmar, M. (2006). *Survival analysis: A practical approach*. J. Wiley.

Mazué, G. P., Dechaume-Moncharmont, F. X., & Godin, J.-G. J. (2015). Boldness–exploration behavioral syndrome: Interfamily variability and repeatability of personality traits in the young of the convict cichlid (*Amatitlania siquia*). *Behavioral Ecology, 26*(3), 900–908.

McCune, K., Blaisdell, A., Johnson-Ulrich, Z., Sevchik, A., Lukas, D., MacPherson, M., Seitz, B., & Logan, C. J. (2023). Using repeatability of performance within and across contexts to validate measures of behavioral flexibility. *PeerJ, 11*, Article e15773.

McCune, K. B., Jablonski, P. G., Lee, S., & Ha, R. R. (2019). Captive jays exhibit reduced problem-solving performance compared to wild conspecifics. *Royal Society Open Science, 6*, Article 181311. https://doi.org/10.1098/rsos.181311

McCune, K. B., Valente, J. J., Jablonski, P. G., Lee, S., & Ha, R. R. (2022). Social behavior mediates the use of social and personal information in wild jays. *Scientific Reports, 12*. Article 2494.

Munda, M., Rotolo, F., & Legrand, C. (2012). parfm: Parametric frailty models in R. *Journal of Statistical Software, 51*(11), 1–20. https://doi.org/10.18637/jss.v051.i11

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface, 14*(134). https://doi.org/10.1098/rsif.2017.0213. Article 20170213.

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews, 85*, 935–956. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Peignier, M., Araya-Ajoy, Y. G., Bégué, L., Chaloupka, S., Dellefont, K., Leeb, C., Walsh, P., Ringler, M., & Ringler, E. (2022). Exploring links between personality traits and their social and non-social environments in wild poison frogs. *Behavioral Ecology and Sociobiology, 76*. https://doi.org/10.1007/s00265-022-03202-9. Article 93.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ramírez, I., Paiva, V. H., Fagundes, I., Menezes, D., Silva, I., Ceia, F. R., Phillips, R. A., & Garthe, S. (2016). Conservation implications of consistent foraging and trophic ecology in a rare petrel species. *Animal Conservation, 19*(2), 139–152.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews, 82*, 291–318. https://doi.org/10.1111/j.1469-185X.2007.00010.x

Schielzeth, H., & Nakagawa, S. (2013). Nested by design: Model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution, 4*, 14–24. https://doi.org/10.1111/j.2041-210x.2012.00251.x

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution, 8*, 1639–1644.

Stroeymeyt, N., Franks, N. R., & Giurfa, M. (2011). Knowledgeable individuals lead collective decisions in ants. *Journal of Experimental Biology, 214*(18), 3046–3054.

Suresh, K., Severn, C., & Ghosh, D. (2022). Survival prediction models: An introduction to discrete-time modeling. *BMC Medical Research Methodology, 22*. https://doi.org/10.1186/s12874-022-01679-6. Article 207.

Takola, E., Krause, E. T., Müller, C., & Schielzeth, H. (2021). Novelty at second glance: A critical appraisal of the novel object paradigm based on meta-analysis. *Animal Behaviour, 180*, 123–142. https://doi.org/10.1016/j.anbehav.2021.07.018

Therneau, T. (2022). Coxme: Mixed effects Cox models. http://ftp.edu.ee/pub/cran/web/packages/coxme/vignettes/coxme.pdf.

Therneau, T. (2024). A package for survival analysis in R. https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf.

Therneau, T., Crowson, C., & Atkinson, E. (2024). Using time dependent covariates and time dependent coefficients in the Cox model. https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf.

Vámos, T. I. F., & Shaw, R. C. (2024). Consistent individual differences give rise to 'caching syndromes' in a food-storing passerine. *Animal Behaviour, 211*, 43–51. https://doi.org/10.1016/j.anbehav.2024.02.012

van den Heuvel, K., Quinn, J. L., Kotrschal, A., & van Oers, K. (2023). Artificial selection for reversal learning reveals limited repeatability and no heritability of cognitive flexibility in great tits (*Parus major*). *Proceedings of the Royal Society B: Biological Sciences, 290*(2003). Article 20231067.

Villagomez, G. N., Nürnberger, F., Requier, F., Schiele, S., & Steffan-Dewenter, I. (2021). Effects of temperature and photoperiod on the seasonal timing of western honey bee colonies and an early spring flowering plant. *Ecology and Evolution, 11*(12), 7834–7849.

Villegas-Ríos, D., Réale, D., Freitas, C., Moland, E., & Olsen, E. M. (2017). Individual level consistency and correlations of fish spatial behaviour assessed from aquatic animal telemetry. *Animal Behaviour, 124*, 83–94.

Waser, N. M., Price, M. V., & Shaw, R. G. (2000). Outbreeding depression varies among cohorts of *Ipomopsis aggregata* planted in nature. *Evolution, 54*(2), 485–491.

Wolf, M., & Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology & Evolution, 27*, 452–461. https://doi.org/10.1016/j.tree.2012.05.001

Wright, T. F., Eberhard, J. R., Hobson, E. A., Avery, M. L., & Russello, M. A. (2010). Behavioral flexibility and species invasions: The adaptive flexibility hypothesis. *Ethology Ecology & Evolution, 22*, 393–404. https://doi.org/10.1080/03949370.2010.505580

Wróbel, A., Kurek, P., Bogdziewicz, M., Dobrowolska, D., & Zwolak, R. (2022). Avian dispersal of an invasive oak is modulated by acorn traits and the presence of a native oak. *Forest Ecology and Management, 505*, Article 119866.

Zwolak, R. (2018). How intraspecific variation in seed-dispersing animals matters for plants. *Biological Reviews, 93*, 897–913. https://doi.org/10.1111/brv.12377