Drivers of age-specific survival in a long-lived seabird: contributions of observed and hidden sources of heterogeneity

Lise M. Aubry[1,2,3]*, **Emmanuelle Cam**[2], **David N. Koons**[3], **Jean-Yves Monnat**[4] **and Samuel Pavard**[5]

[1]*Max Planck Institute for Demographic Research, Rostock, Germany;* [2]*Laboratoire Evolution et Diversité Biologique, Université P. Sabatier, Toulouse, France;* [3]*Department of Wildland Resources and the Berryman Institute, Utah State University, Logan, USA;* [4]*Penn ar Run Izella, Goulien, Cap Sizun, France; and* [5]*Muséum National d'Histoire Naturelle – Centre National de la Recherche Scientifique UMR 14 5145, Paris, France*

Summary

1. We assessed the relative influence of variability in recruitment age, dynamic reproductive investment (time-specific reproductive states) and frailty (unobserved differences in survival abilities across individuals) on survival in the black-legged kittiwake. Furthermore, we examined whether observed variability in survival trajectories was best explained by immediate reproductive investment, cumulative investment, or both.

2. Individuals that delayed recruitment ($\geq$ age 7) suffered a higher mortality risk than early recruits (age 3), especially later in life, suggesting that recruitment age may be an indicator of individual quality. Although recruitment age helped explain variation in survival, time-varying reproductive investment had a more substantial influence.

3. The dichotomy of attempting to breed or not explained variability in survival across life better than other parameterizations of reproductive states such as clutch size, brood size or breeding success. In the kittiwake, the sinequanon condition to initiate reproduction is to hold a nest site, which is considered a very competitive activity. This might explain why attempting to breed is the key level of investment that affects survival, independent of the outcome (failure or success).

4. Interestingly, the more individuals cumulate reproductive attempts over life, the lower their mortality risk, indicating that breeding experience may be a good indicator of parental quality as well. In contrast, attempting to breed at time t increased the risk of mortality between t and $t + 1$. We thus detected an immediate trade-off between attempting to breed and survival in this population; however, the earlier individuals recruited, and the more breeding experience they accumulated, the smaller the cost.

5. Lastly, unobserved heterogeneity across individuals improved model fit more (1·3 times) than fixed and dynamic sources of observed heterogeneity in reproductive investment, demonstrating that it is critical to account for both sources of individual heterogeneity when studying survival trajectories. Only after simultaneously accounting for both sources of heterogeneity were we able to detect the 'cost' of immediate reproductive investment on survival and the 'benefit' of cumulative breeding attempts (experience), a proxy to individual quality.

Key-words: age at first reproduction, Breslow estimator, frailty, individual quality, reproductive investment, senescence, survival analysis, trade-offs

Introduction

Organisms must eventually face trade-offs and allocate limited time and energy among growth, reproduction and survival (Williams 1966). One particular trade-off that has received great attention is that between investment in current reproduction at the expense of future reproduction and (or) survival. Although a large number of experimental and correlative studies have examined this trade-off (Stearns 1992), empirical support for it in wild organisms remains

ambiguous (Harshman & Zera 2007), especially in long-lived species where some seemingly escape trade-offs and appear to be 'Darwinian demons' (i.e. hypothetical organism that can maximize all aspects of fitness simultaneously; Law 1979).

A number of factors might limit our ability to detect trade-offs between reproduction and survival in long-lived species. First, reproductive investment early in life might not bear its effect until much later in life as a result of subtle costs accumulating over time. In long-lived species 'cumulative costs of reproduction' might be the norm rather than the exception (Aubry *et al.* 2009b). Moreover, reproductive traits can be fixed (e.g. age at first reproduction) or can fluctuate over an individual's life in a stochastic manner (changes in egg production, chick production, number of offsprings fledged, etc.) in response to environmental conditions, competition and previous life experiences (Tuljapurkar, Steiner & Orzack 2009). It is thus crucial to account for all of the above sources of variation, whether they are fixed (e.g. age at first reproduction) or dynamic (time or age-varying reproductive states), in order to detect the underlying relationship between survival and reproductive investment.

Second, ecologists often encounter the problem of limited data. High levels of extrinsic mortality in the wild can prevent most individuals from reaching old age, which constitutes the key sample in senescence studies, i.e. a decline in survival at advanced ages (e.g. Ricklefs & Scheuerlein 2001). Thus, long-term monitoring is essential for studying senescence in survival.

Third, a variety of genetic, maternal and environmental factors can lead to variation in survival abilities among individuals of the same population (Wilson & Nussey 2010). When difficult or impossible to measure directly, these unobserved differences in survival abilities across individuals (commonly called 'frailty') lead to underlying changes in the composition of a sample population. According to Vaupel & Yashin's definition (1985), 'frail' individuals readily die and thus exit the sample, leaving only the most 'robust' individuals in the sample at advanced ages. As a result, population-level estimates of age-specific survival can reflect patterns resulting from 'within-generation phenotypic selection' (Endler 1986), rather than genuine age-specific variation in survival experienced by individuals (Vaupel & Yashin 1985; for an application see Fox *et al.* 2006).

The black-legged kittiwake (*Rissa tridactyla*) is a long-lived seabird that exhibits substantial variability in reproductive traits across individuals (Cam & Monnat 2000b; Cam *et al.* 2002a,b; Aubry *et al.* 2009b; Aubry, Cam & Monnat 2009a), some of which are fixed (e.g. the age at first reproduction) and some of which are dynamic over life (age or time-specific reproductive investments). Such sources of observed individual heterogeneity in reproductive investment may help explain variation in survival trajectories, but if so, trade-offs are not readily apparent in kittiwakes. Aubry *et al.* (2009b) found that the cost of early-life reproductive investment on future reproduction was delayed and not born out until late life. Furthermore, age at recruitment, sometimes a proxy to individual quality (Forslund & Pärt 1995), had a large influence on the age trajectory and rate of senescence in breeding success (Aubry *et al.* 2009b). We suspect that any potential trade-off between reproduction and age-specific survival in kittiwakes would also be delayed and tempered by variation in individual quality.

In addition to observed sources of individual heterogeneity, substantial amounts of unobserved heterogeneity have been detected in this population, both in survival and reproduction (Cam & Monnat 2000b, Cam *et al.* 2002a; Aubry *et al.* 2009b). Kittiwakes thus serve as an ideal biological model to evaluate the relative contributions of observed (i.e. reproductive traits) and unobserved individual heterogeneity (i.e. frailty) to variation in survival trajectories and elucidate possible trade-offs between reproductive investment and future survival.

We propose to examine whether (i) trade-offs exist between survival and immediate or cumulative levels of reproductive investment. On one hand, individuals that accumulate several years of reproductive investment might incur long-term somatic costs that could translate into a decline in survival later in life (i.e. senescence in survival). On the other hand, individuals that breed successfully for several consecutive breeding seasons might be of higher intrinsic quality or benefit from breeding experience, which could translate into maintaining high levels of survival throughout life. (ii) While investigating these trade-offs, we account for the potential importance of individual variability in recruitment age on survival trajectories. Recruitment age might be another proxy to individual (parental) quality (Aubry *et al.* 2009b) and temper the trade-offs defined in objective (i). (iii) We implement these effects in parallel with a frailty variable to quantify the relative contributions of observed heterogeneity and frailty to variability in survival trajectories. Doing so may also help elucidate trade-offs between reproduction and survival in a long-lived species that could go undetected otherwise.

Materials and methods

The population of interest has been under intense monitoring for 30 years (2046 individuals, 8279 observations), and all individuals are detected and observed every year from the age at first reproduction until death (here, inferences about mortality are necessarily restricted to the study area; Cam, Cooch & Monnat 2005). Classical survival models used in human demography (e.g. Kleinbaum & Klein 2005) are therefore appropriate for estimating trajectories of survival across life (e.g. Wintrebert *et al.* 2005). Various extensions to the nonparametric Kaplan & Meier (1958) estimator, such as the Cox proportional hazard model (CPH; Cox 1972a) further allow identification of the measurable (i.e. observed) covariates associated with patterns in survival trajectories while accounting for frailty (Klein 1992). Information about the study population and site can be found in Appendix S1. All analyses were conducted in R (Development Core Team 2008; version 2.10.1).

MODELLING OBSERVED HETEROGENEITY

To address objectives (i) and (ii), we used CPH models (library 'survival' in R, procedure 'coxph') that are semi-parametric and have the advantage of making no assumption regarding shape of the underlying mortality hazard (a.k.a., the force of mortality) over life. Each covariate within the model is assumed to act multiplicatively (i.e. proportionally) on the baseline mortality hazard at each time step (e.g. Bradburn *et al.* 2003), such as $h(t, X_i) = h_0(t) \times \exp(\sum_{i=1}^{p} \beta_i X_i)$ where h_0 refers to the baseline hazard (i.e. hazard's value when all covariate values are null), p denotes the number of parameters in the model, the βs denotes a set of estimated parameters, and the Xs represents the data, or series of covariate values for each individual i such as $X = (X_1, X_2, \dots X_i)$, and t denotes time (in our case, time elapsed since recruitment rather than actual age). X_i can either consist of one unique value per individual (e.g. the age at first reproduction) or can be a vector of values (i.e. one value per year lived for each individual; e.g. time-specific reproductive investment).

The study is particular in that individuals that did not attempt to breed at least once were not part of the sample. Individuals enter the 'risk set' at first reproduction (e.g. age 3, 4, 5, 6, 7 or more), which in our study is considered time 0. Thus, time in the above-defined CPH models is a correlate to age and is equivalent to the number of years elapsed since first reproduction. For example, if a bird starts to breed at age 3, time step 1 corresponds to the interval between age 3 and age 4, time step 2 to the interval between age 4 and age 5, for such an individual.

Cox proportional hazard models are widely used to assess the effect of covariates on survival, whereas accelerated failure time models are usually used to assess the underlying form of the mortality hazard, which was of less interest. To test the validity of using CPH models, we used the 'cox.zph' procedure in the 'survival' library of R (Therneau & Grambsch 2000) to assess whether each covariate modality within the best performing CPH model acted proportionally to the mortality hazard. If so, the *P*-value associated with each covariate's proportionality test would be > 0.05. Using interactions between covariates can help relax the assumption of proportional covariate effects, as covariate effects may not only vary across covariate modalities, but can also be tempered by time or age (Martinussen & Scheike 2006).

To examine the relationship between recruitment age, reproductive investment and future survival (objective *i*), we used four biological covariates (see Appendix S2 for graphical representation of these covariates). We used a single time-varying reproductive covariate 'REP' treated as a factor to capture effects of immediate reproductive investment at time t on survival from t to $t + 1$. We first considered 11 different levels of reproductive investment (REP₁) that included not attempting to breed (reproductive level 1), attempting to breed but failing to produce any eggs (level 2); producing one egg (level 3), two eggs (level 4) or three eggs (level 5) but no chick; producing one egg (level 6), two eggs (level 7) or three eggs (level 8) but only fledged a single chick; producing two eggs (level 9) or three eggs (level 10) and successfully fledged two chicks; and producing three eggs that all fledged (level 11). We considered that an individual was attempting to breed if it completed nest construction

Table 1. Cox proportional hazard models testing for the effects of various levels of immediate (i.e. 'REP'; models 1–10), cumulative reproductive investment (i.e. CREP; models 11–20), or both (i.e. REP + LCREP; models 21–120), as well as the effect of interactions (i.e. between REP and LCREP; models 121–220) on age-specific survival. For each model defined below, additive year effects (i.e. 'YEAR') were systematically included to account for environmental variability in survival. Moreover, interactions between the age at first reproduction (i.e. AFR) and CREP and LCREP were considered to account for potential differences in individual quality reflected by the timing of first reproduction. Descriptions of the different levels of reproductive investment for each covariate (REP, CREP and LCREP) are described in the text and outlined below

Model covariates	Models									
Immediate effect of reproductive investment on survival 'REP'	1	2	3	4	5	6	7	8	9	10
Cumulative effect of reproductive investment on survival 'CREP'	11	12	13	14	15	16	17	18	19	20
Immediate and lagged cumulative effects of reproductive investment on survival 'REP + LCREP'					21–120					
Interactions between immediate and lagged cumulative effect of reproduction on survival 'REP * LCREP'					121–220					

Partitioned levels of reproductive investment used to define the models above

		1	2	3	4	5	6	7	8	9	10
$n = 989$	Level 1. does not breed	x	x	x		x		x		x	x
$n = 872$	Level 2. bred but did not lay eggs	x		x		x	x	x	x	x	
$n = 1635$	Level 3. produced 1 egg but no chick	x			x						x
$n = 2068$	Level 4. produced 2 eggs but no chick	x				x	x	x	x	x	
$n = 61$	Level 5. produced 3 eggs but no chick	x									
$n = 638$	Level 6. produced 1 egg, fledged 1 chick	x	x								x
$n = 2004$	Level 7. produced 2 eggs, fledged 1 chick	x		x				x	x	x	
$n = 95$	Level 8. produced 3 eggs, fledged 1 chick	x			x	x	x				
$n = 1200$	Level 9. produced 2 eggs, fledged 2 chicks	x									x
$n = 79$	Level 10. produced 3 eggs, fledged 2 chicks	x						x	x	x	
$n = 27$	Level 11. produced 3 eggs, fledged 3 chicks	x								x	x

n represents the number of observations per level of reproductive investment.

(Maunder & Threlfall 1972), as only individuals that are truly involved in reproduction are able to complete the structure (Cam *et al.* 1998).

We progressively collapsed the different levels of reproductive investment into fewer categories to examine alternative hypotheses regarding the most relevant levels of reproductive investment that affect survival trajectories (Table 1). For example, we distinguished between individuals that failed to breed successfully (level 1 and 2) and individuals that bred successfully (level 3–11) and indexed the covariate as REP_3. Table 1 lists all of the biological subhypotheses tested (models 1–10).

We also considered a cumulative version of REP (i.e. CREP) and investigated the effects of CREP cumulated from the age at first reproduction 'AFR' (i.e., Age at First Reproduction) to time t on survival from time t to $t + 1$ (Table 1; models 11–20). For CREP, we also progressively collapsed the different levels of cumulative reproductive investments and indexed CREP accordingly from $CREP_{11}$ to $CREP_{20}$ (Table 1).

To assess whether both immediate and cumulative reproductive investment affect survival, we developed a lagged cumulative reproductive investment covariate that only included investment from AFR up to time $t − 1$ and its effect on survival from time t to $t + 1$ (and called it 'LCREP'; Appendix S2). This covariate allowed us to examine independent additive effects of cumulative (LCREP, i.e. Lagged Cumulative Reproduction Investment) and immediate (REP) reproductive investment on future survival (i.e. it avoids the qualms of colinearity between covariates). Again, we considered progressively collapsed levels of cumulative reproductive investment and indexed LCREP as mentioned previously. Test of hypotheses pertaining to the effect of immediate (i.e. $REP_{1–10}$), cumulative (i.e. AFR * $CREP_{1–10}$) or both immediate and cumulated reproductive investment ($REP_{1–10}$ + AFR * $LCREP_{1–10}$) on survival from one age to the next was defined in models 21–120 (Table 2).

Finally, we considered a series of models accounting for interactions between REP and LCREP (Table 1; models 121–220), as they could account for the effect of individual differences in reproductive investment on survival better than additive models.

We controlled for differences in AFR across individuals when examining the influence of cumulative reproduction on survival (i.e.

Table 2. Test of hypotheses pertaining to the effect of immediate (i.e. $REP_{1–10}$), cumulative (i.e. AFR * $CREP_{1–10}$) or both cumulated and immediate reproductive investment (AFR + $REP_{1–10}$ * $LCREP_{1–10}$) on survival from one age to the next (i.e. models in Table 2). We only present the top 10 models (out of 220 models) as only the top two models were at all supported by the data (shaded models)

Model	K	AICc	ΔAICc
32	7	16589·66	0·00
132	7	16593·27	3·61
42	7	16612·33	22·67
142	7	16616·57	26·91
62	7	16641·54	51·88
162	7	16645·24	55·58
82	7	16647·04	57·38
102	7	16647·08	57·42
202	7	16650·66	61·00
182	7	16650·67	61·01

K: number of parameters in the model.
AICc, Akaike's information criterion.

by modelling interactions between AFR and CREP or between AFR and LCREP). We also systematically considered an additive effect of year ('YEAR') in each CPH model to account for environmentally driven changes in survival over time.

Overall, we compared the fit of the above-defined CPH models (Table 1) using Akaike's information criterion adjusted for sample size (AICc; Akaike 1973). We based our inference on the top performing model and any model that was within two AICc units of the top model (Burnham & Anderson 2002).

We calculated Pearson's product-moment correlation coefficient for each pair of covariates that were present in the best performing model (Hollander and Wolfe 1973), where $−1 ≤ ρ(X,Y) ≤ 1$ (0 indicates no correlation, 1 a perfect positive correlation and $−1$ a perfect negative correlation).

MODELLING FRAILTY

Vaupel & Yashin (1985) introduced the idea of a frailty component 'z' that acts multiplicatively on the hazard rate at each time step to correct for unobserved individual heterogeneity. Such frailty terms were later implemented into CPH frailty models: $h(t|X_i(t)) = z ×h_0(t) × \exp(\beta_i X_i(t))$, also denoted as $h(t|X_i(t)) = h_0(t) × \exp(\beta_i X_i(t) + \varepsilon)$ with the error term $\varepsilon = \log(z)$ (Klein 1992). The frailty term is generally assumed to have a gamma distribution such that frailty values are positive, whereas a normal distribution ranges from $−∞$ to $+∞$. The expected value of a gamma-distributed frailty for the ith subject exiting at time y_i and experiencing an event of type δ_i is given by: $E(Z|y, \delta, \alpha) = \frac{1+\sigma^2\delta}{1+\sigma^2 H_0(y) \exp^{\beta X}}$ where i is the individual, y is the exit time (i.e. time at which the individual left the study either because it died, or because the study ended), α is the right-censoring indicator ('0' if right-censored, i.e. the individual was still alive at the end of the study; '1' if death is observed), and β is the covariate profile (β_t in the case of a time-varying covariate). We maximized Breslow's maximum likelihood estimator (Breslow 1972) with the E-M algorithm to estimate the variance σ^2 of the frailty term z with mean 1 (Klein 1992; Lin 2007).

We did not use information criterion to compare the mixed model with frailty to marginal models because robust quantification of the appropriate penalty for adding a random effect has not been resolved and is an area of current statistical research (e.g. Jiang *et al.* 2008; Vaida & Blanchard 2005). Instead, we calculated deviance values for a model accounting only for observed heterogeneity (i.e. reproductive covariates; top performing model) (model a), a model accounting for both observed heterogeneity and frailty (model b), and a model accounting for frailty only (i.e. frailty) (model c). Models a, b, and c accounted for a categorical effect of 'YEAR' as a baseline source of environmental variability in survival. A reference model accounting only for temporal changes in survival was also considered (model d; YEAR effect only).

We then used Adler & HillRisLambers' approach (2008), adapted from Zheng (2000) to calculate the relative contribution of observed heterogeneity (i.e. reproductive covariates) and frailty to individual variation in survival. According to this approach, if A is the R^2 of model a, B is the R^2 of model b, and C is the R^2 of model c, then $D = A + C − B$, where D is the overlap of A and B. We were particularly interested in the relative proportion of the variance that is explained only by observed sources of heterogeneity $(A − D)$ and unobserved sources $(C − D)$. However, because the use of R^2 is not appropriate for mixed models with nonlinear link functions, we used Zheng's goodness of fit measure; the percentage reduction in deviance, an appropriate measure for both marginal and mixed models (Zheng 2000). We first calculated deviances from each model's likeli-

hood according to the following formula: $\text{Dev}_x = -2\times \ln(\text{Likelihood}_{\text{model}(x)})$. We then calculated the percentage reduction in deviance according to Zheng's formula: $\text{Dev}_x' = 1 - \frac{\text{Dev}_x}{\text{Dev}_y}$, where Dev stands for deviance, x is the model of interest (either a, b, or c), and y is the reference model (d). We used Dev_x' values calculated for each model (a, b, c and d) and replaced those values in Adler and HillRisLambers' equation to obtain the 'relative percentage reduction in the deviance' attributable to the reproductive covariates on one hand and frailty on the other hand. The ratio of these percentage reductions indicates whether observed reproductive covariates or frailty led to a greater reduction in overall deviance. We note that it is not appropriate to think of Dev_x' as a direct surrogate for R^2 statistics. A 100% reduction in deviance is impossible and thus these percentage reductions should not be thought of on a 0–100% scale (Zheng 2000).

Results

To appropriately account for temporal changes in time-specific survival, we investigated the fit of a fully year-varying CPH model (d.f. = 26), then used the model's estimates to cluster years that had a similar effect on time-specific survival (i.e. similar β estimates). Accordingly, we found that categorizing 'YEAR' into four groups explained annual changes in survival in the most parsimonious fashion (group 1: 2003, 2005, 2006, 2007; group 2: 1983, 1992, 1993, 1996, 1997, 2002, 2004; group 3: 1987, 1988, 1990, 1998, 1999, 2001; group 4: 1984, 1985, 1986, 1989, 1991, 1995, 2000), where group 1 represents the baseline group with the lowest mortality. Naturally, years are not necessarily consecutive within each group because environmental conditions (e.g. storms, massive predation events, pollution, etc.) change stochastically.

Among the set of models testing for trade-offs between reproductive investment and survival, one model was supported by the data (model 32, Table 2). The top model retained the same covariates as the next best performing model and only differed in complexity by only the number of interaction terms. Frailty made an important contribution towards improvement in model fit (large reduction in deviance). Therefore, we present here the estimates associated with model 32 containing an additional frailty term.

The best performing model retained an effect of YEAR treated as a factor, an effect of AFR, REP_2 and LCREP_2, as well as an interaction between AFR and LCREP_2 (see Table 3 for parameter estimates). Positive coefficient estimates for the 'YEAR' effect ($\beta > 0$ or $\exp(\beta) > 1$; Table 3) indicate higher mortality risk and lower survival than the baseline YEAR group (i.e. group 1), and negative parameter coefficients indicate the opposite. Thus, YEAR groups 2, 3 and 4 were years in which mortality risk was higher than in the baseline group (Table 3).

We found that individuals that delay recruitment had lower survival after recruitment than those that began reproduction earlier in life (Table 3, $\exp(\beta)$ = 1·098: 9·8% higher mortality risk per year of delayed age at first reproduction).

The LCREP_2 and REP_2 parameters in the top model indicate that two key levels of reproductive investment accounted for variability in kittiwake survival better than other parameterizations: not attempting to breed (Table 1, reproductive level 1) vs. attempting to breed regardless of clutch size and breeding success (reproductive levels 2–11 treated the same). We found that individuals who accumulated more reproductive attempts up to time $t - 1$ experienced higher subsequent survival from time t to $t + 1$ (Table 3; $\exp(\beta)$ = 0·282). Increasing the number of lagged cumulative breeding attempts by 1 unit decreased the risk of dying by 71·8%. Moreover, the interaction terms including LCREP_2 and AFR indicated that the negative influence of delayed recruitment may be somewhat counterbalanced by the accumulation of breeding attempts over life (Table 3, Fig. 1).

Only after variation in recruitment age (AFR) and breeding experience (LCREP_2) were accounted for, could the 'immediate' impact of reproductive investment on survival be revealed. The effect of REP_2 in the top model indicated that attempting to breed at time t doubled the risk of dying between time t and $t + 1$ (Table 3; $\exp(\beta)$ = 1·963; Fig. 1b,d) relative to individuals that did not attempt to breed (Fig. 1a,c).

Our results further indicated that once individuals cumulated more than 4 lagged breeding attempts ($\text{LCREP}_2 > 4$), their predicted probability of survival over their remaining life was nearly constant and extremely high (Fig. 1), yet only 27·61% of the sample cumulated >4 breeding attempts, mostly young recruits. Thus, the stabilization of the predicted survival surfaces at high survival (Fig. 1) did not apply to many individuals in our population, especially those with delayed recruitment.

Table 3. Coefficient estimates (β) for the top performing model (32) including a frailty term; $\exp(\beta)$ is the associated mortality risk (i.e. a value superior to 1 indicates an increased mortality risk compared to the baseline coefficient, and vice versa, a value inferior to 1 indicates a lower mortality risk). We also provide standard errors [i.e. SE (β)], as well as test statistics

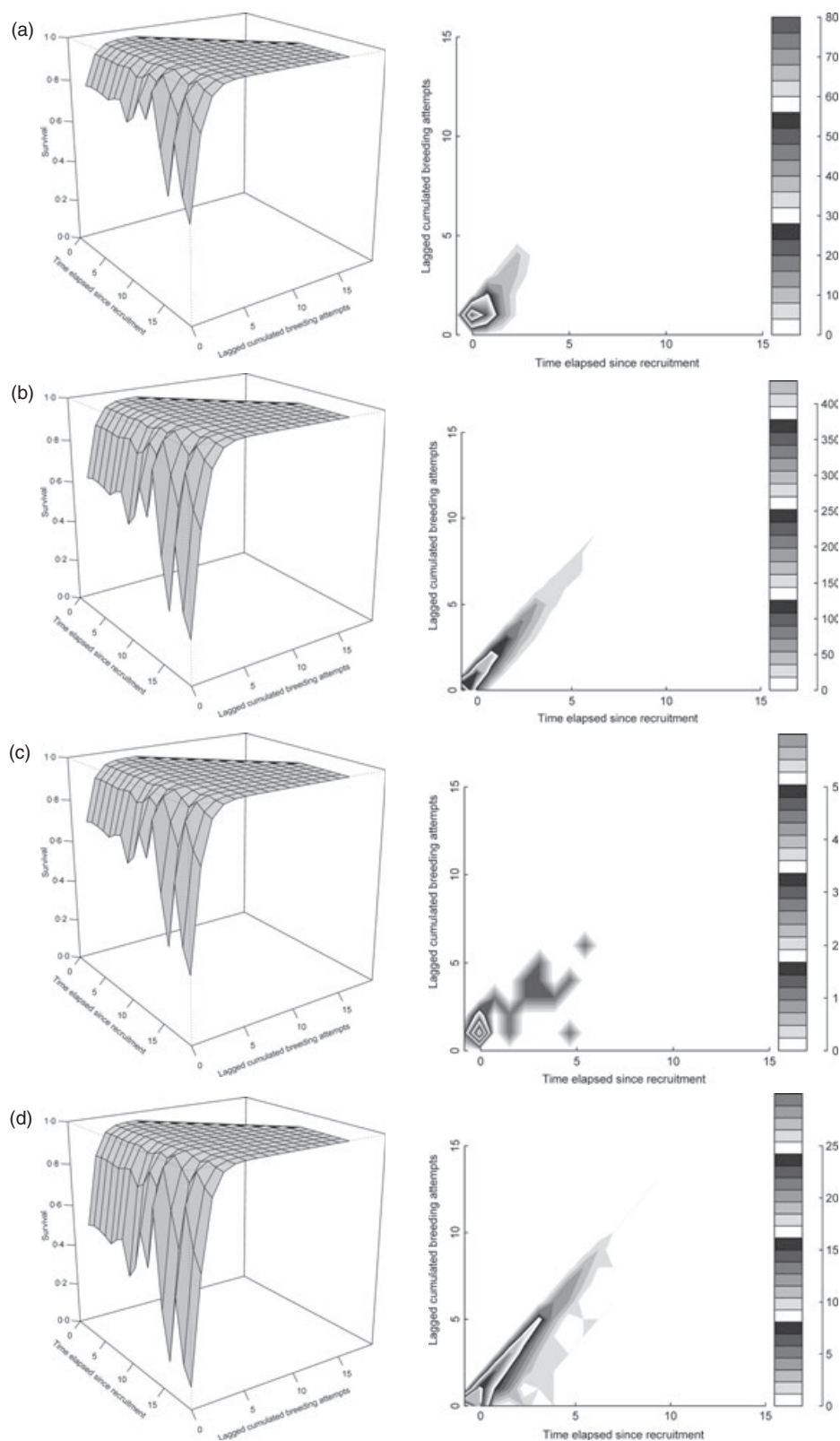| Covariates	β	$\exp(\beta)$	SE (β)	z-values	P-values
YEARC_2	0·264	1·303	0·076	3·5	<0·001
YEARC_3	0·372	1·450	0·082	4·5	<0·001
YEARC_4	0·511	1·666	0·076	6·7	<0·001
AFR	0·094	1·098	0·040	2·3	0·019
REP_2	0·674	1·963	0·087	7·7	<0·001
LCREP_2	−1·266	0·282	0·056	−22·8	<0·001
AFR*LCREP_2	−0·023	0·977	0·012	−1·9	0·057

Fig. 1. Predicted surfaces of survival, averaged across all years, from the best performing Cox Proportional Hazard model accounting for both observed heterogeneity and frailty. Survival is presented across 'time since recruitment' and lagged cumulative breeding attempts (i.e. $LCREP_2$) for various combinations of recruitment age (i.e. AFR) and immediate reproductive investment (i.e. REP_2). We selected four combinations of AFR and REP_2 representing the extremes within each trait: AFR = 3 (panels a & b) and AFR = 7 (panels c & d) (earliest and latest possible recruitment, respectively), and REP_2 = 0 (panels a & c) or 1 (panels b & d) (did not, or did attempt to breed at age x). For each combination of AFR and REP_2, we present contour plots representing sampling size as a function of the time elapsed since recruitment and the number of lagged breeding attempts cumulated over a lifetime ($LCREP_2$).

Table 4. Pearson correlation coefficients indicating the potential direction and strength of the association between pairs of covariates that were present in the best performing model (i.e. $YEAR_C$, AFR, REP_2, and $LCREP_2$)

X	Y	$\rho(X,Y)$
$YEAR_C$	AFR	-0.056
$YEAR_C$	REP_2	0.047
$YEAR_C$	$LCREP_2$	-0.066
AFR	REP_2	0.031
AFR	$LCREP_2$	-0.033
REP_2	$LCREP_2$	0.03

d.f. = 8277.

On the other hand, 72·39% of the population cumulated < 5 breeding attempts over life ($LCREP_2 < 5$), and within this subgroup, most individuals recruited at ages 3, 4, or 5, and cumulated 2 or less breeding attempts overall (Appendix S3). For individuals with < 5 lagged accumulated breeding attempts, survival estimates exhibited substantial variation across reproductive covariate combinations and across life (Fig. 1a–d). For example, an individual that recruited early and did not subsequently breed experienced fluctuations in annual survival between 0·7 and 1·0 throughout much of life (age 3–16), but if alive, experienced rapid senescence to an annual survival probability of 0·31 at age 20 (Fig. 1a). Attempting to breed led to an immediate cost of reproduction on one hand (Fig. 1b), but on the other hand, increased breeding experience (i.e. greater $LCREP_2$), later improved annual chances of survival up to age 16 and reduced the severity of senescence thereafter (and little senescence at all after $LCREP_2 > 4$; Fig. 1a). The same patterns in age-specific survival related to reproductive attempts were also observed in individuals with delayed recruitment (Fig. 1c,d). Attempting to breed and delayed recruitment have similar dramatic effects on survival and the two combined result in the lowest survival chances across life (Fig. 1d).

Of importance, Pearson correlation coefficients indicated no evidence for correlation between covariates retained in the best performing model (Table 4), indicating that all covariates were independent of one another. Residual plots for the best performing model including frailty did not indicate signs of large departure from proportional effects across recruitment groups (AFR; $\rho = 0.013$, $\chi^2 = 0.279$, $P = 0.597$; plots are not presented for the sake of conciseness). The main effect of immediate reproductive investment on the hazard was not proportional (REP_2, $\rho = -0.114$, $\chi^2 = 19.325$, $P < 0.05$), nor was the lagged (main) effect of cumulative reproductive investment ($LCREP_2$, $\rho = 0.363$, $\chi^2 = 213.967$, $P < 0.05$). However, the interaction between AFR and $LCREP_2$ corrected for the occasional non-proportional effect of $LCREP_2$ levels on survival (AFR * $LCREP_2$, $\rho = 0.004$, $\chi^2 = 0.033$, $P = 0.856$), indicating that the proportionality assumption was reasonable as long as the interaction was accounted for.

The estimated variance of the frailty term in the top ranked model that accounted for both sources of heterogeneity was large (3·14). Moreover, we found that unobserved heterogeneity (i.e. frailty) contributed 1·3 times more to the overall reduction in deviance than observed (fixed and dynamic) heterogeneity in reproduction relative to the reference model with only temporal variation in survival (model d). Given that the estimated value of frailty was large and contributed greatly to reduction in model deviance, we felt it was important to include along with observed covariates when making inference about individual variation in trajectories of survival over life.

Discussion

Variation in kittiwake survival was related to both fixed (recruitment age) and dynamic (time-varying reproductive investments and temporal effects) covariates, as well as unobserved individual differences (frailty).

Age at first reproduction had an important impact on survival and senescence. In the black-legged kittiwake, we suspect that early recruits might possess inherent reproductive, survival and competitive abilities that could allow them to start breeding earlier, senesce slower and attain higher fitness than individuals that delay recruitment (e.g. Cam & Monnat 2000a; Cam et al. 2002a; Aubry et al. 2009b). That said, recruiting at the earliest possible age may not be the best strategy either.

Pre-breeding experience can help achieve high levels of reproductive success through 'prospection'. For example, black-legged kittiwakes are known to prospect for breeding sites in order to gain information on reproductive success within a reproductive colony and increase chances of successful settlement and reproduction the following year (e.g. Boulinier et al. 1996). Because chicks are often left unattended by their parents at the nest, 'squatters' (Cadiou 1993) often visit these nests, not to experience parenthood (i.e. squatters often kill the left-alone chicks by beating them, sitting on them, or even re-building nests on them), but to acquire a social status (i.e. squatters become familiar with neighbours whose aggressiveness progressively decreases), while the parents are gone at sea to find food.

Post-recruitment experience on the other hand is the experience gained through previous breeding opportunities. Aubry et al. (2009b) showed that individuals recruiting at intermediate ages maintained high levels of breeding success over their life span, balanced pre- and post-recruitment experience in an advantageous way (i.e. highest levels of breeding success at first reproduction) and seemed to balance the level, onset and speed of reproductive senescence compared to earlier and later recruits (Aubry et al. 2009b). Future work will thus aim at estimating individual fitness (e.g. Coulson et al. 2006) and measure the force of selection on recruitment age operating through its impact on lifetime trajectories of reproductive success (Aubry et al. 2009b) and survival (presented here).

Although recruitment age helped explain variation in survival, dynamic reproductive investment had a more substantial influence. The effects of different levels of reproductive investment (i.e. attempting and not attempting to breed)

on survival demonstrated that dynamic sources of reproductive investment should not be neglected while studying the effects of reproduction on survival (Tuljapurkar, Steiner & Orzack 2009; Steiner, Tuljapurkar & Orzack 2010), especially in long-lived species that have more opportunities to display variation in reproductive investment. We found that both immediate and cumulative reproductive investment influenced kittiwake survival. On one hand, breeding at time t had a large negative effect on survival between time t and $t + 1$ (i.e. the classic immediate cost of reproduction). One the other hand, the more breeding attempts were cumulated up to time $t - 1$, the lower the chance of dying from time t to time $t + 1$. Thus, it seems that immediate costs of reproduction on survival do exist in this population, but in the long run, they could be tempered by the number of cumulated breeding attempts, which may be indicative of environmental experience, innate individual (parental) quality or quality determined by developmental conditions before recruitment. In part, this reinforces the findings of Cam *et al.* (2002a) who observed a positive correlation between breeding probability (analogous to breeding attempts in our study) and survival.

Kittiwakes also experienced significant temporal stochasticity in survival. In recent years (2003–2007), the mortality hazard was lower than in the past (Table 3); however, we know very little about the causal drivers of temporal variation in age-specific or cohort-specific survival. Frederiksen *et al.* (2007) found that kittiwake survival in England and Ireland was negatively correlated with sea-surface temperature and breeding productivity, but positively correlated with an increase in abundance of their principal prey, *Calanus* copepods. Knowledge about changes in resource availability, climatic conditions and the frequency of predation events is accruing for locations near our study area. In the future, we hope to investigate how these variables influence temporal variation in survival for different birth cohorts.

In accordance with Cam *et al.* (2002b), our work also indicates that there is a significant amount of unobserved individual heterogeneity in survival chances (3·14). In fact, our most interesting finding was that frailty (Vaupel & Yashin 1985) reduced the relative model deviance (i.e. improved model fit) 1·3 times more than observed heterogeneity in reproductive investment. The approach usually taken in demographic studies is to try and explain as much variability as possible *via* measured covariates (e.g. Wintrebert *et al.* 2005). Reproductive covariates alone, however, were not sufficient to explain individual variability in adult survival. There is a large amount of individual variation in survival that we cannot explain with the measured covariates, which may be related to genetic differences, micro-habitat variability or traits that we simply did not record. Because it is impossible to capture all of the individual heterogeneity in survival chances with measured (i.e. observed) covariates, we recommend always considering the contribution of unobserved heterogeneity while studying age-specific demographic trajectories (Vaupel & Yashin 1985).

Survival analyses with frailty parameters are rarely used in population and evolutionary ecology (but see for e.g. Fox *et al.* 2006), despite their growing popularity in human demography (Hougaard 1991). In large part, this is likely driven by the requirement of perfect detectability, a condition that has recently been relaxed in capture-mark-recapture methods (e.g. Royle 2008; Gimenez & Choquet 2010). Moreover, identifiability of frailty in CPH models can potentially be confounded with a lack of proportionality (K. Wachter, pers. com.). Given that the observed covariates in the top model generally satisfied the assumption of proportionality, we do not think this was of great significance in our study, but careful attention should be paid to this issue until better statistical methods are developed. Development of mixed models accounting for fixed and dynamic covariates, as well as dynamic frailty, is also needed (Tuljapurkar, Steiner & Orzack 2009). Such models might be able to explain even more variability in age-specific reproduction and survival for long-lived species and could thus be critical in efforts to learn more about long-term trade-offs in the wild.

Additional questions regarding frailty and key life-history traits involved in trade-offs remain. For example, how heritable is an individual's risk of mortality? In human demography, correlated gamma-frailty models of bivariate survival in pairs of twins are used to decompose frailty into genetic and environmental components, allowing for estimation of heritability in frailty (e.g. Iachine *et al.* 1998). In wild animal populations, animal models are also used to understand evolutionary mechanisms underlying variation in key life-history traits (Kruuk 2004). Quantitative genetics is providing a fertile research framework to understand the evolution of life-histories that we intend to use in future research on kittiwakes (e.g. Hadfield 2010; Papaïx *et al.* 2010).

References

Adler, P.B. & HilleRisLambers, J. (2008) The influence of climate and species composition on the population dynamics of ten prairie forbs. *Ecology*, **89**, 3049–3060.

Akaike, H. (1973) *Information Theory and an Extension of the Maximum Likelihood Principle*. International Symposium on Information Theory (eds B.N., Petran & D.F., Csaki), pp. 267–281. Akadéniai Kiadi, Budapest, Hungary.

Aubry, L.M., Cam, E. & Monnat, J.-Y. (2009a) Habitat selection, age-specific recruitment, and reproductive success in a long-lived seabird, the black-legged kittiwake. *Modeling Demographic Processes in Marked Populations Series* (eds D.L. Thomson, E.G. Cooch & M.J. Conroy). *Environmental and Ecological Statistics*, **3**, 365–392.

Aubry, L.M., Koons, D.N., Monnat, J.-Y. & Cam, E. (2009b) Consequences of recruitment decisions and heterogeneity on age-specific breeding success in a long-lived seabird. *Ecology*, **90**, 2491–2505.

Boulinier, T., Danchin, E., Monnat, J.-Y., Doutrelant, C. & Cadiou, B. (1996) Timing of prospecting and the value of information in a colonial breeding bird. *Journal of Avian Biology*, **27**, 252–256.

Bradburn, M.G., Clark, T.G., Love, S.B. & Altman, D.G. (2003) Survival analysis part II: multivariate data analysis: an introduction to concepts and methods. *British Journal of Cancer*, **89**, 431–436.

Breslow, N.E. (1972) Discussion of the paper by D. R. Cox. *Journal of the Royal Statistical Society B*, **34**, 216–217.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. Springer-Verlag, New York, USA.

Cam, E., Cooch, E.G. & Monnat, J.-Y. (2005) Earlier recruitment or earlier death? On the assumption of homogeneous survival in recruitment studies *Ecological Monographs*, **75**, 419–434.

Cam, E. & Monnat, J.-Y. (2000a) Apparent inferiority in first time breeders in the kittiwake: the role of heterogeneity among age-classes. *Journal of Animal Ecology*, **69**, 380–394.

Cam, E. & Monnat, J.-Y. (2000b) Stratification based on reproductive success reveals contrasting patterns of age-related variation in demographic parameters in the kittiwake. *Oikos*, **90**, 560–574.

Cam, E., Hines, J.E., Monnat, J.-Y., Nichols, J.D. & Danchin, E. (1998) Are adult nonbreeders prudent parents? The Kittiwake model. *Ecology*, **79**, 2917–2930.

Cam, E., Cadiou, B., Hines, J.E. & Monnat, J.-Y. (2002a) Influence of behavioural tactics on recruitment and reproductive trajectory in the kittiwake. *Journal of Applied Statistics*, **291**, 163–185.

Cam, E., Link, W.A., Cooch, E.G., Monnat, J.-Y. & Danchin, E. (2002b) Individual covariation between life-history traits: seeing the trees despite the forest. *American Naturalist*, **159**, 96–105.

Coulson, T., Benton, T.G., Lundberg, P., Dall, S.R.X., Kendall, B.E. & Gaillard, J.M. (2006) Estimating individual contributions to population growth: evolutionary fitness in ecological time. *Proceedings of the Royal Society of London Biology series*, **273**, 547–555.

Cox, D.R. (1972a) Regression models and life-tables. *Journal of the Royal Statistical Society B*, **34**, 187–220.

Endler, J.A. (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, USA.

Forslund, P. & Pärt, T. (1995) Age and reproduction in birds-hypotheses and tests. *Trends in Ecology and Evolution*, **10**, 374–378.

Fox, G.A., Kendall, B.E., Fitzpatrick, J.W. & Woolfenden, G.E. (2006) Consequences of heterogeneity in survival probability in a population of Florida scrub-jags. *Journal of Animal Ecology*, **75**, 921–927.

Frederiksen, M., Edwards, M., Mavor, R.A. & Wanless, S. (2007) Regional and annual variation in black-legged kittiwake breeding productivity is related to sea surface temperature. *Marine Ecological Progress Series*, **350**, 137–143.

Gimenez, O. & Choquet, R. (2010) Individual heterogeneity in studies on marked animals using numerical integration: capture–recapture mixed models. *Ecology*, **91**, 148–154.

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

Harshman, L.G. & Zera, A.J. (2007) The cost of reproduction: the devil in the details. *Trends in Ecology and Evolution*, **222**, 80–86.

Hollander, M. & Wolfe, D.A. (1973) *Nonparametric Statistical Methods*. John Wiley & Sons, New York, USA.

Hougaard, P. (1991) Modelling heterogeneity in survival data. *Journal of Applied Probabilities*, **28**, 695–701.

Iachine, I.A., Holm, N., Harris, J.R., Begun, A.Z., Iachina, M.K., Laitinen, M., Kaprio, J. & Yashin, A. (1998) How heritable is individual susceptibility to death? The results of an analysis of survival data on Danish, Swedish and Finnish twins *Twin Research*, **1**, 196–205.

Jiang, J., Rao, S.J., Gu, Z. & Nguyen, T. (2008) Fence methods for mixed model selection. *Annals of Statistics*, **36**, 1669–1692.

Kaplan, E.L. & Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.

Klein, J.P. (1992) Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics*, **48**, 795–806.

Kleinbaum, D.G. & Klein, M. (2005) *Survival Analysis: A Self-Learning Text*, 2nd edn. Springer-Verlag, New York, USA.

Kruuk, L. (2004) Estimating genetic parameters in natural populations using the 'animal model'. *Proceedings of the Royal Society of London Biology series*, **359**, 873–890.

Law, R. (1979) Optimal life histories under age-specific predation. *American Naturalist*, **114**, 399–417.

Lin, D.Y. (2007) On the Breslow estimator. *Lifetime Data Analysis*, **13**, 471–480.

Martinussen, T. & Scheike, T.H. (2006) *Dynamic Regression Models for Survival Data*, Springer ed, New York, USA.

Maunder, J.E. & Threlfall, W. (1972) The breeding biology of the black-legged kittiwake in Newfoundland. *Auk*, **89**, 789–816.

Papaïx, J., Cubaynes, S., Buoro, M., Charmantier, A., Perret, P. & Gimenez, O. (2010) Combining capture-recapture and pedigree data to estimate heritability of survival. *Journal of Evolutionary Biology*, **23**, 2176–2184.

R., Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R. Foundation for statistical computing, Vienna, Austria.

Ricklefs, R.E. & Scheuerlein, A. (2001) Comparison of aging-related mortality among birds and mammals. *Experimental Gerontology*, **36**, 845–857.

Royle, J.A. (2008) Modeling individual effects in the Cormack-Jolly-Seber Model: a state-space formulation. *Biometrics*, **64**, 364–370.

Stearns, S.C. (1992) *The Evolution of Life Histories*, 2nd edn. Oxford University Press, Oxford, England.

Steiner, U.K., Tuljapurkar, S. & Orzack, S.H. (2010) Dynamic heterogeneity and life history variability in the Kittiwake. *Journal of Animal Ecology*, **79**, 436–444.

Therneau, T. & Grambsch, P. (2000) *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York, USA.

Tuljapurkar, S., Steiner, U.K. & Orzack, S. (2009) Dynamic heterogeneity in life histories. *Ecology Letters*, **11**, 1–14.

Vaida, F. & Blanchard, S. (2005) Conditional Akaike information for mixed effects models. *Biometrika*, **92**, 351–370.

Vaupel, J.W. & Yashin, A.I. (1985) Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *American Statistician*, **39**, 176–185.

Williams, G.C. (1966) *Adaptation and Natural Selection*. Princeton University Press, Princeton, USA.

Wilson, A.J. & Nussey, D.H. (2010) What is individual quality? An evolutionary perspective *Trends in Ecology and Evolution*, **1197**, 1–8.

Wintrebert, C.M.A., Zwinderman, A.H., Cam, E., Pradel, R. & van Houwelingen, J.C. (2005) Joint modeling of breeding and survival in the kittiwake using frailty models. *Ecological Modelling*, **181**, 203–213.

Zheng, B. (2000) Summarizing the goodness of fit of generalized linear models for longitudinal data. *Statistical Medicine*, **19**, 1265–1275.

Supporting Information

The following Supporting Information is available for this article online.

Appendix S1. Study site and sample specifications.

Appendix S2. Representation of the four reproductive covariates of interest: AFR.

Appendix S3. Percentage of individuals within the population recruiting at ages 3, 4, 5, 6, 7 or more, and cumulating 1, 2, 3, 4, or more than 5 breeding attempts over life (CREP = 1, 2, 3, 4, $\geq$5).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.