Dealing with death data: individual hazards, mortality and bias

Michael S. Zens and David R. Peart

6044 Gilman Hall, Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA

In ecology and evolution, we have barely begun to tap the information available in survival data. Who lives or dies, and why, is a large part of natural selection. In ecology, these are key questions for building better individual-based models of population and community dynamics. Powerful analytical tools exist to answer them, but the literature is scattered across disciplines, and its relevance is often obscured by inconsistent terminology and technical presentation. Here, we evaluate methods for the application of such tools to ecology and evolution. Analyses based on individual hazards of death are particularly promising, especially in combination with improvements in sampling design. The same methods can also reduce the largely unrecognized biases that plague population-level estimates of mortality rates.

Estimates of fitness, including survival, are fundamental to evolutionary research. In ecology, individual-based approaches have deepened our understanding of population and community processes [1,2]. It is extraordinary, then, that death rates have historically been treated mostly at the population level. By contrast, the other fundamental fitness component, reproductive success, is routinely quantified for individuals. This anomaly probably arises from a lack of familiarity with individual-based approaches to analyzing death rates. The concept of HAZARD (see Glossary), which is central to the methods that we discuss here, defines the probability that an individual will die during a time interval. The death or survival of any one individual over a time period tells us little about its hazard function. However, effective methods do exist to estimate individual hazards given sufficient death event data (Table 1). They have been applied extensively in other disciplines [3–5] but less commonly in ecology and evolution, where studies of mark–recapture in birds [6–9] and of senescence [10–14] are among the best examples.

Mortality rate estimators are more familiar and are easily calculated from the number of deaths in a sample of known initial size. They are intended to estimate the per-capita death rate, aggregated over individuals and over a sample period. However, they tend to blind us to the substantial individual hazard variation that exists in natural and experimental populations [4,13,15]. They have also produced systematic biases in population-level mortality estimates [15].

Here, we present a framework for analyzing death rates when hazards vary, both among individuals and over time. We classify statistical methods by their potential to represent individual differences and temporal changes in hazard. We recommend analytical methods and sampling principles to: (a) estimate environmental and genetic effects on individual hazard; (b) quantify hazard variability; and (c) reduce bias in population mortality rates.

Individual hazard, variability and population mortality rate

Individuals clearly differ in attributes that influence hazard. Genetic variation in life span and death rate occurs in laboratory populations (e.g. fruit flies *Drosophila melanogaster* [16,17], medflies, *Ceritis capitata* [10], and nematodes *Caenorhabditis elegans* [18]) and field populations (e.g. water fleas *Daphnia* spp. [19], salmon *Oncorhynchus kisutch* [20], and cottonwood trees *Populus trichocarpa* [21]). Individual differences in environmental experience, both past and present, also strongly affect hazard. Documented examples include shading of plants [22], density of snowshoe hares *Lepus americanus* in different cover types [23], and cooperative feeding and nest defense in scrub jays *Aphelocoma coerulescens* [7]. Environmental experience can interact with genetic differences to amplify variation in hazard [24,25]. In addition, the effects of both genetic and environmental factors on hazard can change as individuals age [26,27].

If hazards are identical among individuals, the MEAN HAZARD for individuals in a sample is the same as the mortality rate calculated from the same data [28–30]. But even when extreme measures are taken in the laboratory [11,31,32], it is difficult to engineer situations in which individual hazards actually are equal. When the rigorous application of a mortality rate estimator depends on the demanding and unlikely assumptions of homogeneity, we refer to it as a 'restricted estimator of mortality' (REM).

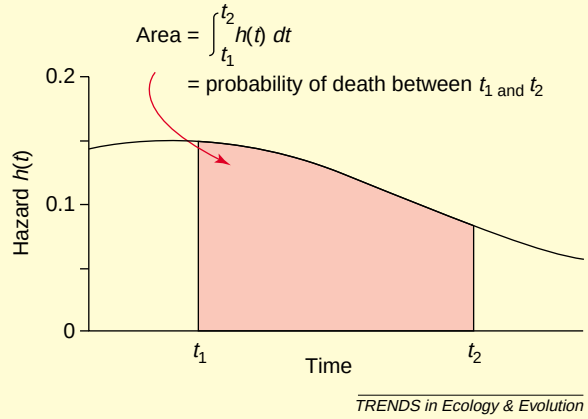
Here, we argue that REMs, such as the commonly used RESTRICTED MORTALITY RATE, $\hat{\lambda}(t)$, always underestimate actual death rates, and suggest ways to reduce this bias. Efforts to control the bias are so rare in our field that little information is currently available to quantify bias. We supplement a brief review of evidence and analytical methods with hypothetical examples (Box 1, Fig. 1), to illustrate relationships among death-rate parameters.

Corresponding author: Michael S. Zens (zens@dartmouth.edu).

http://tree.trends.com 0169-5347/03/$ - see front matter doi:10.1016/S0169-5347(03)00096-X

Glossary

Hazard $h(t)$: a function describing temporal change in the instantaneous death rate experienced by individuals in a sample. Commonly referred to as the 'force of mortality' or the 'mortality density.' More precisely, hazard is the probability density function that generates the probability of dying in a time interval. Units: number of deaths individual-at-risk^{-1} time^{-1}. Hazards and hazard analysis can be applied to events other than deaths.



Hazard, baseline $\overline{h}_0(t)$: the mean of instantaneous hazards in a sample of N individuals, if the initial hazard distribution of that sample is maintained (i.e. replenished as deaths occur). Provides an unbiased estimate of mean instantaneous death rate in a heterogeneous population. Units: number of deaths individual-at-risk^{-1} t^{-1} (Eqn I).

$$\overline{h}_0(t) = \frac{1}{N_0}\sum_{i=0}^{N_0} h_i(t) \qquad \text{[Eqn I]}$$

Hazard, mean $\overline{h}(t)$: the mean instantaneous hazard in a sample of N individuals. In heterogeneous samples, reflects both changes in individual hazards through time, and the change in population composition resulting from the disproportionate loss of high hazard individuals. Units: number of deaths individual-at-risk^{-1} t^{-1} (Eqn II). For a broad range of hazard distributions (represented by a range of α values in Eqn III), $\overline{h}(t)$ underestimates baseline hazard, $\overline{h}_0(t)$, more when $\overline{h}_0(t)$ is greater, hazard variance, σ^2, is larger, and time duration, t, is longer [34].

$$\overline{h}(t) = \frac{1}{N_t}\sum_{j=1}^{N_t} h_j(t) \qquad \text{[Eqn II]}$$

$$\overline{h}(t) = \frac{\overline{h}_0(t)}{\left[1+\alpha^{-1}\sigma^2\int_0^t \overline{h}_0(u)du\right]^{\alpha}} \quad [34] \qquad \text{[Eqn III]}$$

Mortality rate, baseline $\overline{H}_0(t)$: hazard averaged across individuals and the sample period t. Provides an unbiased estimate of sample death rate over a time period. Units: number of deaths individual-at-risk^{-1} t^{-1} (Eqn IV).

$$\overline{H}_0(t) = \frac{1}{t}\int_0^t \overline{h}_0(u)du \qquad \text{[Eqn IV]}$$

Mortality rate, restricted $\hat{\lambda}(t)$: hazard averaged across individuals and the sample period t, assuming individual hazards are identical. Systematically underestimates the mean death rate when hazards are heterogeneous (Box 1 and Fig. 1). Units: number of deaths individual-at-risk^{-1} t^{-1}. $\hat{\lambda}(t)$ is related to hazards by (Eqn V).

$$\hat{\lambda}(t) = \frac{1}{t}\int_0^t \overline{h}(u)du \qquad \text{[Eqn V]}$$

Commonly calculated as (Eqn VI):

$$\hat{\lambda}(t) = -\log_e[N_t/N_0]/t \qquad \text{[Eqn VI]}$$

Box 1. Heterogeneity and bias: dynamics of demographic samples in a hypothetical experiment

A two-year experiment is run to test the effect of fungus exposure on seedling mortality. The population comprises two equally abundant phenotypes, susceptible (S-type) and immune (I-type). This hazard heterogeneity is invisible to the investigator and hazards are constant in time. Randomly chosen individuals are assigned to fungus application or control. Control individuals (both S- and I-types) have a hazard of 0.10 deaths individual-at-risk^{-1} month^{-1}, as do I-types in the treatment. The fungus raises the hazard for S-types to 0.30 (Fig. 1a, main text, illustrates the treatment group).

Because early deaths in the treatment group are concentrated in the S-types, the hazard distribution changes. The mean instantaneous hazard of survivors, $\overline{h}(t)$, declines, approaching the hazard of the I-types. The temporal trend in $\hat{\lambda}(t)$ is an artifact, because of the declining $\overline{h}(t)$. This cause of the false trend points to a target value for $\hat{\lambda}(t)$ that can be used to calculate bias. Using baseline hazard, $\overline{h}_0(t)$ (the mean hazard expected if the sample maintained the initial hazard distribution), we define a baseline, time-averaged mortality rate, $\overline{H}_0(t)$, which is constant at 0.2, the mean for a sample that maintains S-types (hazard 0.3) and I-types (0.1) at equal abundance*.

$\hat{\lambda}(t)$ for the homogeneous control remains constant at 0.1. $\hat{\lambda}(t)$ for the treatment is 0.172 after six months, 0.151 after 12 months and 0.129 after two years. Consequently, we underestimate the fungus effect (true value 0.2 − 0.1 = 0.1) by 28% (six months), 49% (12 months) or 71% (two years). Neither individual hazards nor the effect of the fungus actually change over time.

Individuals are always at risk of death, so we emphasize continuous time death rates [e.g. $\overline{H}_0(t)$, $\hat{\lambda}(t)$] but patterns are similar for discrete time estimates. The proportion dead, (for treatment, control pairs) are (0.64, 0.45), (0.84, 0.70) and (0.95, 0.91) at six, 12 and 24 months, respectively. Estimates for discrete monthly mortality estimates, $m(t)$, (treatment, control pairs) are (0.158, 0.095), (0.140, 0.095) and (0.121, 0.095), with an unbiased treatment death rate of 0.181 and biases in the fungus effect of 24%, 46% and 69%, respectively. For each of $\hat{\lambda}(t)$, proportion dead and $m(t)$, the estimated treatment effect declines with interval length.

* The curves for $\overline{h}_0(t)$ and $\overline{H}_0(t)$ coincide only when individual hazards are constant in time.

Bias is inevitable unless hazard heterogeneity is incorporated in the analysis

To estimate a mortality rate, deaths must be sampled over a period. Yet, once deaths occur, the hazard composition of a heterogeneous sample changes (Box 1, Fig. 1), so that the sample begins to differ from the target population in precisely the measure of interest (the mean death rate experienced by individuals). We assume here that the researcher is interested in estimating the mortality rate for a population represented by the initial demographic sample (i.e. with its initial hazard distribution). Bias in $\hat{\lambda}(t)$ is inevitable for any heterogeneous sample, regardless of the method used to estimate mortality (e.g. life tables or survival analysis). But bias can be reduced if hazard heterogeneity is recognized and incorporated into the analysis.

The impact of hazard heterogeneity on estimates of death rates has a long history of analysis in medicine [4,33], industrial quality control [3,34], human demography [28,35,36] and econometrics [5,37,38]. The causes and patterns of hazard variability differ between animals and air conditioners, but differences in hazard, and the consequent generation of bias, are universal. In ecology and evolution, in spite of some recent attention [6,8,9,14,15,25,39,40], the problem of bias in mortality rates is neither widely recognized nor commonly addressed.

Three factors determine the magnitude of bias in REMs from heterogeneous populations. First, high initial variability in hazard increases bias, because the persistent

Table 1. Estimators of death rates, cross-classified by their capacity to deal with heterogeneous and time-varying hazards

			Hazard changes in time			
	1. Hazard constant in time	**Refs**	**2. Defined in an *a priori* function[a]**	**Refs**	**3. Estimated from event data[b]**	**Refs**
Hazard heterogeneity						
A. Identical hazards	**A1:**		**A2:**		**A3:**	
	Life table methods	[77,78]	Life table methods	[77,78,87]	Life table methods	[88]
	Matrix model	[54,79]	Mark–recapture survival	[85,86]	Mark–recapture survival	[86]
	Difference equation, e.g. $m(t)$	[79,80]	Time-dependent survival	[30]	Piecewise hazard	[70]
	Mayfield method	[81,82]	analysis		Mayfield method:	
	Differential equation, e.g. $\lambda(t)$	[79,83,84]			modified	[89,90]
	Mark–recapture survival	[85,86]			iterative	[70]
B. Individual differences defined by measured covariates[c] (parametric or nonparametric)	**B1:**		**B2:**		**B3:**	
	Survival analysis:		Time-dependent survival		Life table methods	[78]
	Accelerated failure time	[30,91]	analysis:		Logistic regression[d]	[69,92]
	Proportional hazards	[30]	Accelerated failure time	[30,91]	Piecewise hazard models	[69]
	Logistic regression[d]	[70,92]	Proportional hazards	[30]	Cox regression[d]	[30]
	Life table	[54,71]	Logistic regression[d]	[70,92]	Other survival estimators	[37,65]
	Mark–recapture survival	[6,52]	Mark–recapture survival	[6,85]		
C. Defined in an *a priori* function[a]	**C1:**		**C2:**		**C3:**	
	Frailty models:		Frailty models of hazards	[33]	NPMLE[e] parametric in individual differences but not in time	[5,37,65]
	Hazard analysis	[42,93]	Mark–recapture survival	[6,51]		
	Logistic regression	[92]				
	Mark–recapture survival	[6,51]			Correlated frailty	[94]
D. Estimated from event data[b]	**D1:**		**D2:**		**D3:**	
	NPMLE[e] nonparametric in individual differences	[37,65]	NPMLE[e] parametric in time but not in individual differences	[37,65]	Estimators are unstable[f]	[37,65,95]
	Mover–Stayer models	[74]	Latent random effects models	[9]		
	Latent random effects models	[9]				

[a]Previous data and/or theory can provide a basis for parametric assumptions about the shape of the hazard distribution (heterogeneity) or changes in the baseline hazard through time. With the correct model form, these methods are statistically powerful.
[b]In the absence of explicit theory to predict the form of either the distribution or temporal change of hazards, nonparametric methods estimate these attributes directly from event data.
[c]These methods can associate identifiable or measurable traits with individual hazards.
[d]Logistic regression and Cox models are generally used to estimate relative risks, but can be effective for estimating absolute death rates if combined with other information on time variation in hazards.
[e]NPMLE, Nonparametric maximum likelihood estimator.
[f]When used to estimate both changes in time and individual hazards simultaneously, currently available nonparametric estimators are unreliable and confounded.

individuals (which tend to have low hazards) then differ more from those that die early (and tend to have high hazards). A longer sample period increases bias, because there is more time for the sample to become dominated by the lower hazard individuals. Finally, a high overall death rate increases bias, because then the hazard distribution changes more rapidly. In Box 1, only the sample period varies, but the effects of all three factors are documented elsewhere [33,41] and play an important role in general models of death rates in heterogeneous populations [28,33,38,42–46].

Changing hazards over time

As with individual variation, temporal variation in hazard is difficult to eliminate, even in the laboratory. In the field, weather, seasonality, climate change, migration and disturbance cause environments to change. In the laboratory, resource availability [31] or waste materials [11] can change hazards over time. Even in a constant environment, death rates can vary as individuals age [12,26,39] and/or reproduce [47].

Combined individual and temporal variation can be challenging. Consider the simple hypothetical demographic samples in Fig. 1b and Fig. 1c. Each consists of two subpopulations that have different hazards. Within subpopulations, hazards are identical. Hazards might increase for both subpopulations, as in Fig. 1b, if conditions deteriorate (e.g. due to increasing pollution load, climate change or increasing densities of an

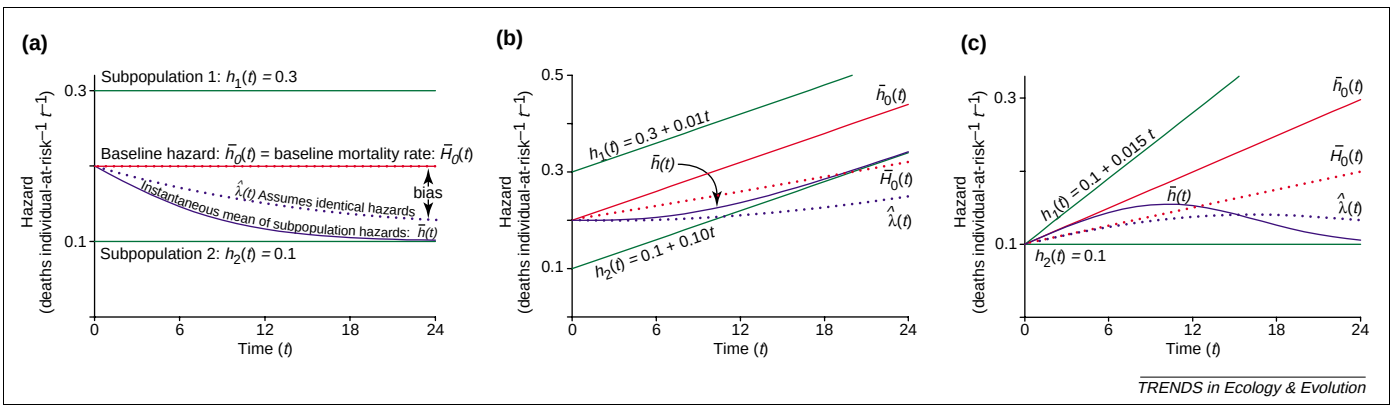


Fig. 1. Heterogeneous and time-varying hazards. Death rates are shown for three hypothetical demographic samples in which individuals are followed through time. In each case, hazards are heterogeneous; the two subpopulations (1 and 2) differ in hazard, but hazards are identical within subpopulations. In (a), hazards are constant in time. In (b), all individuals have linearly increasing hazards, whereas in (c), hazards increase in the first subpopulation but remain constant in the second. Six death rates are shown: the hazards of each subpopulation (h_1 and h_2); the 'baseline' hazard, $\bar{h}_0(t)$ = the hypothetical mean hazard if the hazard composition of the initial sample were maintained; the instantaneous mean hazard of the survivors in the sample, $\bar{h}(t)$; the commonly estimated population mortality rate estimate, $\hat{\lambda}(t)$; and the unbiased baseline mortality rate, $\bar{H}_0(t)$, defined as the time averaged value of $\bar{h}_0(t)$ over the interval $0 - t$. All death rates have the same units and are ultimately derived from individual hazards, so are plotted under the general label 'hazard'. Mortality rate bias is the difference between $\bar{H}_0(t)$ and $\hat{\lambda}(t)$. In these examples, enough time elapses to show full development of trends and biases. In addition to changes in hazards through time, trends in $\hat{\lambda}(t)$ depend on three fundamental quantities: hazard heterogeneity, magnitude of hazard and length of the census interval. Consequences of more complex heterogeneity have been explored elsewhere [41].

introduced competitor or pathogen). Alternatively, adverse changes might affect only one susceptible subpopulation, as in Fig. 1c; this pattern has been suggested, for example, for inciting stresses in some tree populations [48].

In both Fig. 1b and Fig. 1c, the BASELINE HAZARD, $\bar{h}_0(t)$, reflects the time trend in mean death rate for a sample that maintains the original hazard distribution. The trajectory of the REM, $\hat{\lambda}(t)$, reflects neither the trends in individual hazard, nor the unbiased, time-averaged BASELINE MORTALITY RATE, $\bar{H}_0(t)$. (See Box 1 and Glossary for relationships between death rates.) Odd trends can occur in REMs; $\hat{\lambda}(t)$ declines with increasing sample duration in Fig. 1c, even though the hazards in the population are either increasing (subpopulation 1) or constant (subpopulation 2). Other hypothetical examples are illustrated by Vaupel and Yashin [41]. Real examples that combine the effects of heterogeneity and time varying hazards can be more complex (e.g. in the research on senescence described below).

When an average death rate over individuals is needed (e.g. for simulations with population models), we suggest that $\bar{h}_0(t)$ might be more logical and practical than a mortality rate [i.e. $\hat{\lambda}(t)$ or $\bar{H}_0(t)$] that is time-averaged over what might be an arbitrary interval. In fact, we present $\bar{H}_0(t)$ mainly as a construct to evaluate the inherent limitations of $\hat{\lambda}(t)$ as a death-rate parameter (Box 1).

Evidence of mortality bias depends on analysis of hazard heterogeneity

Unlike some performance measures (e.g. growth), death rate cannot be measured directly on individuals, so bias is evident empirically only when a less-biased estimate is available for comparison. Evidence of mortality bias in ecology and evolution is limited, because of the paucity of cases where hazard variability and its causes have been investigated. The most studied examples are those of senescence in cohorts. With advancing age, estimated population death rates plateau in humans *Homo sapiens*, *D. melanogaster* and *C. elegans*, or even decline in *C. capitata*, wasps *Diachasmimorpha longiacaudtis* or yeast *Saccharomyces cerevisiae* [12,13,39]. Investigators asked: do individual hazards really decline with age, or are increases in hazard masked by changes in sample structure (i.e. increasing proportion of low-hazard individuals)? Although neither hypothesis has yet been definitively rejected, the research has advanced the modeling of hazard variability and techniques for the control of hazard variation [12,49,50].

We summarize three cases where hazard heterogeneity was analyzed effectively for field populations of birds. In an exceptionally thorough analysis on kittiwakes *Rissa tridactyla* [8,9,40], declines in annual survival, from 0.84 at four years of age to 0.76 at 16 years, were detected only when individual hazard variation was modeled explicitly.

In a 25-year study of Florida scrub jays *Aphelocoma coerulescens* [7], birds engaging in cooperative feeding and/or nest defense had higher hazards. Dramatic increases in death rates with age (from 0.15 to 0.32 deaths individual^{-1} yr^{-1} over the eight-year reproductive period) were apparent only when this hazard variation was incorporated.

Investigators were apparently successful in removing hazard heterogeneity (and bias) in a 30-year record for mallard ducks *Anas platyrhynchos* [6,51], by stratifying their samples by location, age and sex. After assessing other measured sources of hazard variation, they concluded that residual mortality bias was probably low (≈ 0.05 yr^{-1}, an underestimate of $\sim 10\%$). Finally, studies of waterfowl [52,53] and trees [15] showed clear potential for bias by simulating hazard variation based on other sources of information.

In summary, although documentation of heterogeneity-induced bias is meager, bias is logically inevitable. The practical and theoretical significance of this bias extends to derived rates that are sensitive to estimates of mortality rate (e.g. population growth [54], community dynamics [55], rates of biomass turnover [56] and carbon sequestration [57]).

Experimental and sampling designs to reduce mortality bias

Clever experimental designs can reduce biases associated with heterogeneity by reducing the three contributing factors described (the degree of hazard variability in the sample, the length of the sample interval and the overall death rate). Reducing variability in the sample might not be appropriate if the goal is to estimate a mean mortality rate for a heterogeneous target population. But if homogeneity is desired, hazard variability can be reduced by selecting individuals that are similar in traits that affect (or plausibly affect) hazard (e.g. age, sex, size, condition, genotype and the environmental conditions experienced).

Shortening the census interval can reduce bias, because, when only a small proportion dies, the hazard distribution cannot change much. However, shortening the census interval must generally be compensated by a larger initial demographic sample to obtain sufficient death events for analysis. Of the three factors, death rate might be the least under the researcher's control, although there might be scope for adjustment in some manipulative experiments. As with census interval, there is a tradeoff involving sample size; low death rates demand a larger initial sample [58].

It is important that all three factors be recognized in attempts to avoid bias. For example, it has been suggested that comparing intervals of equal duration using repeat censuses in trees [15,59] can avoid bias. But hazard variability and overall death rates also contribute to heterogeneity bias. Comparisons of mortality rates between heterogeneous populations might be unbiased even if the estimate for each population is biased [43]. However, unbiased comparisons would require equal hazard variability and equal sample duration. They would also require that the very parameters being compared (overall death rates) be similar.

Other specialized approaches can be used. Experimental trials with twins or other matched pairs can reduce individual variability [60–62]. In an ingenious approach, mortality rates in samples have been deliberately boosted in ancillary experiments, accelerating change in the hazard distribution [63,64]. Initial variability in the sample was then assessed, providing a basis for less-biased estimates of death rate in the main experiments.

Finally, even when sampling is designed to reduce bias, it is desirable, where possible, to use analytical techniques based on individual hazards. This further reduces bias in mortality rates and provides information about the factors influencing individual hazards.

Analyses to quantify hazard variation and reduce mortality bias

Consistent violation of assumptions and consequent biases are unsatisfactory and call for an alternative approach to estimating mortality rates. Estimators of individual hazard (Box 2) are an obvious choice to incorporate individual variation in hazards [9,33,37,65–68]. Differences in hazard can be associated with measurable attributes of individuals ('covariates'; Boxes 2,3); Table 1). For example, we might know or suspect that individuals survive better if they are larger, carry particular mutations, or have higher fat reserves. By incorporating covariates into a statistical model, we can estimate their effects on hazard. This enables us to account for changing hazard distributions and to reduce bias that would occur if we were blind to heterogeneity in the sample (Box 3).

In practice, models and analyses might not be entirely individual-based. Qualitative covariates (e.g. helpers versus nonhelpers in the scrub jay study [7]) lead to estimates of mortality rates for subgroups in a population. But even then, it might be possible to cross-classify individuals by several covariates, characterizing individual hazards more effectively.

Changes in hazard through time can be estimated by using models (e.g. using individual hazards) that explicitly incorporate temporal change (Table 1, columns 2 and 3). Hazard estimators have been developed to reduce their dependence on the assumed form of temporal change [69,70], and even to incorporate temporal variation in covariates that influence hazard [30,71,72].

Several methods (Boxes 2,3; Table 1) can make use of covariates to handle both individual and temporal

Box 2. Choosing methods for hazard estimation

Death rates are estimated from the timing of death events, which are usually generated by a continuous underlying process (although there can be pulses of mortality). Therefore, continuous time estimates are generally appropriate [e.g. $h(t)$]. However, discrete time estimates can be made (e.g. using life tables, logistic regression or difference equations).

Even when the timing of individual deaths is imprecisely known, as with interval-censored data*, methods are available for both continuous and discrete time estimates [96]. Shorter intervals produce more precise data about the timing of death events, providing access to more powerful methods for estimating the effects of measured factors (which we call covariates) on death rates.

* An individual is censored if its time of death is unknown. Individuals that survive for an unknown period beyond a sample date are right censored. Similarly, left censored individuals died at some unknown time before a sample date. Typical interval sampling generates both right and left censored individuals ('interval-censored' data). Methods for estimation of death rates require appropriate assumptions about the temporal distribution of censored events [96].

Table 1 (main text) cross-classifies methods according to their capacity to represent temporal change and individual differences in hazard. Methods in cell A1, although commonly used, represent neither. Those in row B were designed mainly to assess the effects of covariates on death rates, but can also be effective to estimate the distribution of hazards among individuals. Methods in rows C and D incorporate hazard heterogeneity using prior theory (row C) or patterns in the timing of death events (row D); they demand precision in the timing of death events and large samples (especially those in row D). Thus, with appropriate data, one can estimate temporal changes in hazard or even the distribution of individual hazards, solely from the timing of death events. However, event data alone are insufficient to estimate simultaneously both temporal trends in hazards and the hazard heterogeneity among individuals (cell D3). Methods in columns 2 and 3 (especially rows A and B) are used to estimate temporal changes in hazard.

Readily accessible sources [30,71,72,96,97] can help those unfamiliar with hazard analysis to match analytical methods to particular situations.

Box 3. Estimating individual hazards, and reducing bias in mean hazard

Although hazards cannot be measured directly on individuals, analysis of factors that contribute to (or are associated with) individual hazard provides an indirect approach. This is useful whether the primary goal is to understand individual hazard variation or to estimate mean hazard (mortality rate) accurately. Consider a hypothetical monitored population of snowshoe hares, whose hazard is increased by their individually measured parasite loads (Fig. Ia, blue line, Fig. Ib). These data can then be combined to estimate the distribution of hazards (owing to parasite load) in Fig. Ic (histogram, blue line) as well as the mean hazard (Fig. Ic, blue arrow). If deaths were monitored, but parasite loads were unknown, individual hazards would be estimated as identical (Fig. Ia, red line, Fig. Ic, red arrow and line). The estimated mean hazard would then be less than that estimated with knowledge of parasite load, even though the observed death events were the same.

Estimation of a less-biased mean death rate, using individual hazards, depends crucially on two factors. First, estimates of individual hazard depend on the form of the statistical model fit to data (as in Fig. Ia). In the absence of previous data or theory, there can be many plausible models, especially when multiple factors are involved, so full use of available techniques to discriminate among alternative models [98] is recommended.

Second, some factors influencing hazard will remain unidentified or unmeasured, or their effects will be inadequately modeled. In the hare example, genetic variation, access to resources, exposure to predators, and previous injury or infection are some of the many possible contributing factors. In the ideal (but impractical) case where all effects are known (Fig. Ic; dashed grey histogram), one could estimate the true mean hazard (Fig. Ic; dashed grey arrow). Clearly, we generally underestimate both hazard variation and the true mean hazard.

The true, underlying hazard distribution can also be approached using other methods [9,37,65] that require high temporal precision in death event data and large sample size, and/or *a priori* theory about the true hazard distribution (Table 1, main text). These do not attempt to expose the sources of individual hazard variation, but they can be combined with covariate analyses that do. Available data constrain analytical options. But even a single strong covariate effect, as in this example, can improve our understanding of individual hazard variation and substantially reduce bias in estimates of population death rate.

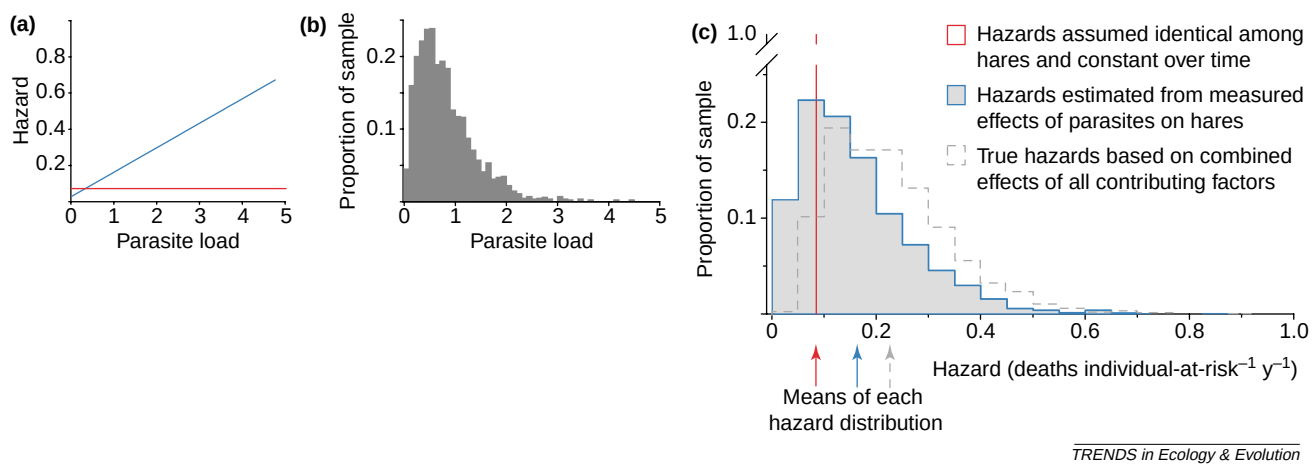


Fig. I.

variation in hazard. In other words, with adequate measures of appropriate individual attributes and adequate samples of death events through time, one can estimate instantaneous individual hazards [30,73], even for heterogeneous populations in changing conditions. Estimates of individual hazard have extraordinary value for estimating individual fitness [9,66], for elucidating the ecological significance of individual variation, and for modeling population and community dynamics in individual-based models [73]. They can also serve as the basis for less-biased population-level estimates [9,66,68].

Analytical methods that include the effects of measured covariates provide powerful techniques to incorporate hazard variability. However, the resulting estimates of hazard are actually predictions of a statistical model (similar to predictions from a regression model). Consequently, they suffer two inherent limitations.

First, they provide no direct measure of total hazard variability (Box 3). We can estimate the influences on hazard only for those covariate factors that we incorporate explicitly [44]. New techniques (Table 1, cells C3, D1 and D2) can lead to glimpses of the total variation in restricted cases, but the influences of unmeasured factors will remain unknown [65,68] (Box 3). Even for the covariates we include, using simplified functions, we might represent their effects inadequately. Thus, we generally underestimate total hazard variability, and therefore underestimate both population mortality rate and the bias that results when variability is ignored [68].

Second, estimates of individual differences in hazard are derived from analyses of aggregate death event data and thus lack independence. Specifically, the estimate (prediction) of hazard for an individual depends not only on its own 'intrinsic' hazard, but also on the hazards of other individuals in the population (just as a prediction from a regression equation depends on all the data analyzed, across the range of the independent variable). Lack of independence can cause underestimates of errors even if the mortality estimates are unbiased. Nevertheless, as hazard models are developed to incorporate more of the underlying causes of individual variability, as well as temporal change, estimates of hazard variability will become more inclusive and mortality estimates less biased.

Individual variability can also be estimated without covariates, but only under rather demanding conditions

(Table 1, rows C and D). Unknown variation can be modeled parametrically assuming individual random effects, if the distribution is assumed [6,39,58,14] (Table 1, cells C2 and C3). Estimates are sensitive to the form of the assumed (often arbitrary) distribution, but recent extensions to random–variable, mixed-model methods might enable tests of the assumptions with field data [9,66,67]. Another alternative is to use nonparametric hazard estimators (Table 1, cell D1) that identify relatively homogeneous subgroups based on similar times-to-death [74]. Finally, in the most sophisticated approach, nonparametric maximum likelihood estimators (NPMLEs; Table 1, cells D1, D2 and C3) assign individuals to an optimal number of relatively homogeneous groups [5,37,65].

Generalizations to other kinds of event data

In ecology and evolution, we analyze a variety of non-death event data, and the concept of hazard and the methods of hazard analysis are applicable to all of them. For example, heterogeneity introduces bias in the estimates of population-level rates for tree recruitment [15], copepod moulting [75], and fire disturbance [76]. Such biases are due to the same processes as those discussed for death rates, and can be reduced by applying similar analytical approaches.

Conclusion

Estimates of death rate in ecology and evolution are commonly biased. Systematic underestimation can produce errors in analysis and prediction, for a variety of practically and theoretically important applications based on estimates of death rates. Collection of covariate data, more effective analytical methods and improvements in sampling design can facilitate estimates of individual hazards and total hazard variability. At the same time, they can reduce bias in estimates of population-level mortality rate. We suggest that wider use of individual hazard models might be the most powerful approach for analyzing death rates. The capacity to associate traits with hazard at the level of the individual has enormous potential benefits for the study of natural selection and evolution.

Acknowledgements

This work was supported by a grant from the Andrew W. Mellon Foundation, by Gilman and Graduate Fellowships from Dartmouth College, and by the British Columbia Ministry of Forests, Forest Practices Branch. James Dykes, Arthur G. Blundell and T. Scott Sillett provided M.S.Z. with formative discussions on missing variable problems in statistics. We thank Arthur G. Blundell, Christina M. Robinson, Carol L. Folt, Matthew P. Ayres and Douglas G. Sprugel for insightful comments on the article.

References

1 McGraw, J.B. and Caswell, H. (1996) Estimation of individual fitness from life-history data. *Am. Nat.* 147, 47–64
2 Pascual, M. and Levin, S.A. (1999) From individuals to population densities: Searching for the intermediate scale of nontrivial determinism. *Ecology* 80, 2225
3 Follmann, D.A. and Goldberg, M.S. (1988) Distinguishing heterogeneity from decreasing hazard rates. *Technometrics* 30, 389–396
4 Aalen, O.O. (2000) Medical statistics – no time for complacency. *Stat. Methods Med. Res.* 9, 31–40
5 Land, K.C. *et al.* (2001) Discrete-time hazard regression models with hidden heterogeneity – the semiparametric mixed Poisson regression approach. *Soc. Methods Res.* 29, 342–373
6 Rexstad, E.A. and Anderson, D.R. (1992) Heterogeneous survival rates of mallards (*Anas platyrhynchos*). *Can. J. Zool. Rev. Can. Zool.* 70, 1878–1885
7 McDonald, D.B. *et al.* (1996) Actuarial senescence and demographic heterogeneity in the Florida scrub jay. *Ecology* 77, 2373–2381
8 Cam, E. and Monnat, J.Y. (2000) Stratification based on reproductive state reveals contrasting patterns of age-related variation in demographic parameters in the kittiwake. *Oikos* 90, 560–574
9 Cam, E. *et al.* (2002) Individual covariation in life-history traits: seeing the trees despite the forest. *Am. Nat.* 159, 96–105
10 Carey, J.R. *et al.* (1992) Slowing of mortality rates at older ages in large medfly cohorts. *Science* 258, 457–461
11 Curtsinger, J.W. *et al.* (1994) Rates of mortality in populations of *Caenorhabditis elegans*. *Science* 266, 826–826
12 Pletcher, S.D. and Curtsinger, J.W. (1998) Mortality plateaus and the evolution of senescence: why are old-age mortality rates so low? *Evolution* 52, 454–464
13 Vaupel, J.W. *et al.* (1998) Biodemographic trajectories of longevity. *Science* 280, 855–860
14 Service, P.M. (2000) Heterogeneity in individual mortality risk and its importance for evolutionary studies of senescence. *Am. Nat.* 156, 1–13
15 Sheil, D. and May, R.M. (1996) Mortality and recruitment rate evaluations in heterogeneous tropical forests. *J. Ecol.* 84, 91–100
16 Curtsinger, J.W. *et al.* (1992) Demography of genotypes – failure of the limited life-span paradigm in *Drosophila melanogaster*. *Science* 258, 461–463
17 Hughes, K.A. and Charlesworth, B. (1994) A genetic-analysis of senescence in *Drosophila*. *Nature* 367, 64–66
18 Brooks, A. *et al.* (1994) Mortality-rates in a genetically heterogeneous population of *Caenorhabditis elegans*. *Science* 263, 668–671
19 Dudycha, J.L. and Tessier, A.J. (1999) Natural genetic variation of life span, reproduction, and juvenile growth in *Daphnia*. *Evolution* 53, 1744–1756
20 Hard, J.J. *et al.* (2000) Genetic variation in mortality of chinook salmon during a bloom of the marine alga *Heterosigma akaskiwo*. *J. Fish Biol.* 56, 1387–1397
21 Dunlap, J.M. *et al.* (1994) Genetic-variation and productivity of *Populus-trichocarpa* and its hybrids.7. 2-year survival and growth of native black cottonwood clones from 4 river valleys in Washington. *Can. J. For. Res.* 24, 1539–1549
22 George, L.O. and Bazzaz, F.A. (1999) The fern understory as an ecological filter: growth and survival of canopy-tree seedlings. *Ecology* 80, 846–856
23 Boonstra, R. *et al.* (1998) The impact of predator-induced stress on the snowshoe hare cycle. *Ecol. Monogr.* 68, 371–394
24 Gaillard, J.M. *et al.* (2000) Temporal variation in fitness components and population dynamics of large herbivores. *Annu. Rev. Ecol. Syst.* 31, 367–393
25 Pletcher, S.D. and Curtsinger, J.W. (2000) The influence of environmentally induced heterogeneity on age-specific genetic variance for mortality rates. *Genet. Res.* 75, 321–329
26 Promislow, D.E.L. *et al.* (1996) Age-specific patterns of genetic variance in *Drosophila melanogaster*.1. Mortality. *Genetics* 143, 839–848
27 Promislow, D.E.L. and Tatar, M. (1998) Mutation and senescence: where genetics and demography meet. *Genetica* 103, 299–314
28 Vaupel, J.W. *et al.* (1979) Impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16, 439–454
29 Vaupel, J.W. and Yashin, A.I. (1983) *The Deviant Dynamics of Death in Heterogeneous Populations*, International Institute for Applied Systems Analysis
30 Cox, D.R. and Oakes, D. (1984) *Analysis of Survival Data*, Chapman & Hall
31 Khazaeli, A.A. *et al.* (1996) Effect of density on age-specific mortality in *Drosophila*: a density supplementation experiment. *Genetica* 98, 21–31
32 Khazaeli, A.A. *et al.* (1998) The fractionation experiment: reducing heterogeneity to investigate age-specific mortality in *Drosophila*. *Mech. Ageing Dev.* 105, 301–317
33 Aalen, O.O. (1988) Heterogeneity in survival analysis. *Stat. Med.* 7, 1121–1137
34 Proschan, F. (1963) Theoretical explanation of observed decreasing failure rate. *Technometrics* 5, 375–383
35 Keyfitz, N. and Littman, G. (1979) Mortality in a heterogeneous population. *Popul. Stud. J. Demogr.* 33, 333–342

36 Gavrilov, L.A. and Gavrilova, N.S. (1991) *The Biology of Life Span: A Quantitative Approach*, Harwood Academic Publishers
37 Heckman, J.J. and Singer, B. (1985) Social science duration analysis. In *Longitudinal Analysis of Labor Market Data* (Heckman, J.J. and Singer, B., eds) pp. 39–110, Cambridge University Press
38 Hsieh, F. (2001) On heteroscedastic hazards regression models: theory and application. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63, 63–79
39 Service, P.M. *et al.* (1998) Experimental evolution of senescence: an analysis using a 'heterogeneity' mortality model. *Evolution* 52, 1844–1850
40 Cam, E. and Monnat, J.Y. (2000) Apparent inferiority of first-time breeders in the kittiwake: the role of heterogeneity among age classes. *J. Anim. Ecol.* 69, 380–394
41 Vaupel, J.W. and Yashin, A.I. (1985) Heterogeneity ruses – some surprising effects of selection on population-dynamics. *Am. Stat.* 39, 176–185
42 Manton, K.G. *et al.* (1981) Methods for comparing the mortality experience of heterogeneous populations. *Demography* 18, 389–410
43 Gail, M.H. *et al.* (1985) Biased estimates of treatment effects in randomized experiments. *Biometrics* 41, 327
44 Hougaard, P. (1991) Modeling heterogeneity in survival data. *J. Appl. Prob.* 28, 695–701
45 Manton, K.G. *et al.* (1997) Analyses of cohort mortality incorporating observed and unobserved risk factors. *Math. Comput. Model.* 25, 89–107
46 Yashin, A.I. and Manton, K.G. (1997) Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. *Stat. Sci.* 12, 20–34
47 Stearns, S.C. (1992) *The Evolution of Life Histories*, Oxford University Press
48 Pedersen, B.S. (1998) The role of stress in the mortality of Midwestern oaks as indicated by growth prior to death. *Ecology* 79, 79
49 Drapeau, M.D. *et al.* (2000) Testing the heterogeneity theory of late-life mortality plateaus by using cohorts of *Drosophila melanogaster*. *Exp. Gerontol.* 35, 71–84
50 Rose, M.R. and Mueller, L.D. (2000) Ageing and immortality. *Philos. Trans. R. Soc. Lond. Ser. B* 355, 1657–1662
51 Burnham, K.P. and Rexstad, E.A. (1993) Modeling heterogeneity in survival rates of banded waterfowl. *Biometrics* 49, 1194–1208
52 Nichols, J.D. *et al.* (1982) Additional comments on the assumption of homogeneous survival rates in modern bird banding estimation models. *J. Wildl. Manage.* 46, 953–962
53 Pollock, K.H. and Raveling, D.G. (1982) Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. *J. Wildl. Manage.* 46, 88–98
54 Caswell, H. (2001) *Matrix Population Models: Construction, Analysis and Interpretation*, Sinauer Associates
55 Pacala, S.W. *et al.* (1996) Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecol. Monogr.* 66, 1
56 Phillips, O.L. *et al.* (1998) Changes in the carbon balance of tropical forests: evidence from long-term plots. *Science* 282, 439–442
57 Hurtt, G.C. *et al.* (1998) Terrestrial models and global change: challenges for the future. *Glob. Change Biol.* 4, 581–590
58 Pletcher, S.D. (1999) Model fitting and hypothesis testing for age-specific mortality data. *J. Evol. Biol.* 12, 430–439
59 Williamson, G.B. *et al.* (2000) Amazonian tree mortality during the 1997 El Nino drought. *Conserv. Biol.* 14, 1538–1542
60 Marenberg, M.E. *et al.* (1994) Genetic susceptibility to death from coronary heart disease in a study of twins. *N. Engl. J. Med.* 330, 1041–1046
61 Yashin, A.I. and Iachine, I.A. (1995) Genetic analysis of durations – correlated frailty model applied to survival of Danish twins. *Genet. Epidemiol.* 12, 529–538
62 Redondo, M.J. *et al.* (2001) Heterogeneity of type Idiabetes: analysis of monozygotic twins in Great Britain and the United States. *Diabetologia* 44, 354–362
63 Khazaeli, A.A. *et al.* (1995) Stress experiments as a means of investigating age-specific mortality in *Drosophila-melanogaster*. *Exp. Gerontol.* 30, 177–184
64 Khazaeli, A.A. *et al.* (1997) Heat-induced longevity extension in *Drosophila*.1. Heat treatment, mortality, and thermotolerance. *J. Gerontol. A Biol. Sci. Med. Sci.* 52, B48–B52
65 Baker, M. and Melino, A. (2000) Duration dependence and nonparametric heterogeneity: a Monte Carlo study. *J. Econ.* 96, 357–393
66 Cooch, E.G. *et al.* (2002) Occam's shadow: levels of analysis in evolutionary ecology – where to next? *J. Appl. Stat.* 29, 19–48
67 Link, W.A. *et al.* (2002) Model-based estimation of individual fitness. *J. Appl. Stat.* 29, 207–224
68 Nichols, J.D. (2002) Discussion comments on: 'Occam's shadow: levels of analysis in evolutionary ecology – where to next?' by Cooch, Cam and Link. *J. Appl. Stat.* 29, 49–52
69 Allison, P.D. (1982) Discrete-time methods for the analysis of event histories. In *Sociological Methodology* (Leinhardt, S., ed.), pp. 61–98, Jossey-Bass
70 Manly, B.F.J. and Schmutz, J.A. (2001) Estimation of brood and nest survival: comparative methods in the presence of heterogeneity. *J. Wildl. Manage.* 65, 258–270
71 Allison, P.D. (1995) *Survival Analysis using the SAS System: A Practical Guide*, SAS Institute
72 Hosmer, D.W. and Lemeshow, S. (1999) *Applied Survival Analysis: Regression Modeling of Time to Event Data*, John Wiley & Sons
73 Kendall, B.E. and Fox, G.A. (2002) Variation among individuals and reduced demographic stochasticity. *Conserv. Biol.* 16, 109–116
74 Blumen, I., *et al.* eds (1955) *The Industrial Mobility of Labor as a Probability Process* Cornell University
75 Souissi, S. and Ban, S. (2001) The consequences of individual variability in moulting probability and the aggregation of stages for modelling copepod population dynamics. *J. Plankton Res.* 23, 1279–1296
76 Polakow, D.A. and Dunne, T.T. (1999) Modelling fire-return interval t: stochasticity and censoring in the two-parameter Weibull model. *Ecol. Model.* 121, 79–102
77 Promislow, D.E.L. (1991) Senescence in natural-populations of mammals – a comparative study. *Evolution* 45, 1869–1887
78 Carey, J.R. (1993) *Applied Demography for Biologists with Special Emphasis on Insects*, Oxford University Press
79 Charlesworth, B. (1994) *Evolution in Age-Structured Populations*, Cambridge University Press
80 Primack, R.B. *et al.* (1985) Growth rates and population structure of Moraceae trees in Sarawak, East Malaysia. *Ecology* 66, 577–588
81 Mayfield, H. (1975) Suggestions for calculating nest success. *Wilson Bull.* 87, 456–466
82 Bart, J. and Robson, D.S. (1982) Estimating survivorship when the subjects are visited periodically. *Ecology* 63, 1078–1090
83 Swaine, M.D. and Lieberman, D. (1987) Special issue: the dynamics of tree populations in tropical forest– 7th Aberdeen Hull Symposium: introduction. *J. Trop. Ecol.* 3, 289–290
84 Sheil, D. *et al.* (1995) The interpretation and misinterpretation of mortality-rate measures. *J. Ecol.* 83, 331–333
85 Pollock, K.H. *et al.* (1990) Statistical inference for capture-recapture experiments. *Wildl. Monogr.* 1, 1–97
86 Pollock, K.H. *et al.* (1995) A capture–recapture survival analysis model for radio-tagged animals. *J. Appl. Stat.* 22, 661–672
87 Gaillard, J.M. *et al.* (1994) Senescence in natural populations of mammals – a reanalysis. *Evolution* 48, 509–516
88 Sibly, R.M. *et al.* (1997) Mortality rates of mammals. *J. Zool.* 243, 1–12
89 Johnson, D.H. (1979) Estimating nest success – Mayfield method and an alternative. *AUK* 96, 651–661
90 Klett, A.T. and Johnson, D.H. (1982) Variability in nest survival rates and implications to nesting studies. *AUK* 99, 77–87
91 Kalbfleisch, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*, John Wiley & Sons
92 Allison, P.D. (1987) Introducing a disturbance into logit and probit regression models. *Soc. Methods Res.* 15, 355–374
93 Hougaard, P. (1984) Life table methods for heterogeneous populations – distributions describing the heterogeneity. *Biometrika* 71, 75–83
94 Yashin, A.I. and Iachine, I.A. (1997) How frailty models can be used for evaluating longevity limits: taking advantage of an interdisciplinary approach. *Demography* 34, 31–48
95 Heckman, J.J. *et al.* (1990) Testing the mixture of exponentials hypothesis and estimating the mixing distribution by the method of moments. *J. Am. Stat. Assoc.* 85, 582–589
96 Lindsey, J.C. and Ryan, L.M. (1998) Tutorial in biostatistics – methods for interval-censored data. *Stat. Med.* 17, 219–238
97 Fox, G.A. (2001) Failure time analysis: studying times-to-events and rates at which events occur. In *Design and Analysis of Ecological Experiments*, 2nd edn, (Scheiner, S.M. and Gurevitch, J., eds) pp. 253–289, Oxford University Press
98 Burnham, K.P. and Anderson, D.R. (1998) *Model Selection and Inference: A Practical Information-Theoretic Approach*, Springer