Ecology, 87(9), 2006, pp. 2356–2365

THE FITTING OF GENERAL FORCE-OF-INFECTION MODELS TO WILDLIFE DISEASE PREVALENCE DATA

DENNIS M. HEISEY,[1,4] DAMIEN O. JOLY,[2] AND FRANÇOIS MESSIER[3]

[1]*USGS-National Wildlife Health Center, Madison, Wisconsin 53711 USA*
[2]*Field Veterinary Program, Wildlife Conservation Society, Bronx, New York 10460 USA*
[3]*Department of Biology, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5E2 Canada*

Abstract. Researchers and wildlife managers increasingly find themselves in situations where they must deal with infectious wildlife diseases such as chronic wasting disease, brucellosis, tuberculosis, and West Nile virus. Managers are often charged with designing and implementing control strategies, and researchers often seek to determine factors that influence and control the disease process. All of these activities require the ability to measure some indication of a disease's foothold in a population and evaluate factors affecting that foothold. The most common type of data available to managers and researchers is apparent prevalence data. Apparent disease prevalence, the proportion of animals in a sample that are positive for the disease, might seem like a natural measure of disease's foothold, but several properties, in particular, its dependency on age structure and the biasing effects of disease-associated mortality, make it less than ideal. In quantitative epidemiology, the "force of infection," or infection hazard, is generally the preferred parameter for measuring a disease's foothold, and it can be viewed as the most appropriate way to "adjust" apparent prevalence for age structure. The typical ecology curriculum includes little exposure to quantitative epidemiological concepts such as cumulative incidence, apparent prevalence, and the force of infection. The goal of this paper is to present these basic epidemiological concepts and resulting models in an ecological context and to illustrate how they can be applied to understand and address basic epidemiological questions. We demonstrate a practical approach to solving the heretofore intractable problem of fitting general force-of-infection models to wildlife prevalence data using a generalized regression approach. We apply the procedures to *Mycobacterium bovis* (bovine tuberculosis) prevalence in bison (*Bison bison*) in Wood Buffalo National Park, Canada, and demonstrate strong age dependency in the force of infection as well as an increased mortality hazard in positive animals.

Key words: *apparent prevalence;* Bison bison; *bovine tuberculosis; force of infection; infection hazard; infection-specific mortality hazard; left truncation; logistic regression;* Mycobacterium bovis; *proportional hazards models; wood buffalo.*

INTRODUCTION

Cumulative incidence vs. apparent prevalence: the force of infection

Increasingly, researchers and wildlife managers find themselves in situations where containment, control, and eventual elimination of infectious wildlife diseases are central goals. Well-publicized examples include chronic wasting disease, tuberculosis, brucellosis, and West Nile virus. Inevitably, in dealing with this disease, it becomes necessary to have some quantitative measure of a disease's foothold in a population. Typical questions that are asked are whether a disease's foothold is changing, and what factors are associated with the strength of the foothold. The most common type of data collected from disease systems in free-ranging animals is apparent prevalence, defined as the proportion of animals in a sample that test positive for the disease of interest. At first consideration, apparent prevalence might seem like a suitable measure of the disease's foothold. Indeed, in many wildlife disease situations, apparent prevalence has been the focus of analysis and evaluation. However, as we will consider in greater detail, apparent prevalence is usually a less-than-ideal comparative measure because of its strong dependency on age structure of the sample. The goal of this paper is to present a general consideration of infectious disease dynamics and to show how this leads to a general approach for analyzing apparent prevalence data that naturally adjusts for age structure and allows for the testing of sharply focused epidemiological hypotheses. The basic problem that we address is estimating the rate at which negative animals become positive: the force of infection. The force of infection is a key quantity in its own right, and it forms the foundation for the much more challenging problem of estimating the transmission rate, the rate at which positive animals infect negative animals (Farrington et al. 2001).

Manuscript received 1 June 2005; revised 19 December 2005; accepted 4 January 2006; final version received 27 February 2006. Corresponding Editor: K. Wilson.

[4] E-mail: dheisey@usgs.gov

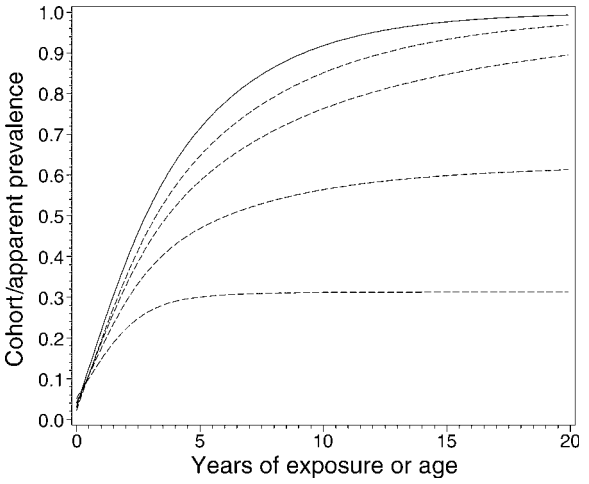

FIG. 1. The solid line is the exposure-specific cumulative incidence $\pi(x)$ for infection hazard (force of infection) $\lambda = 0.25$. The dashed lines are the age-specific apparent prevalence $\nu_c(t)$ for varying values of the infection-associated mortality hazard increment μ (0.1, 0.2, 0.4, and 0.8); smaller μ values correspond to higher $\nu_c(t)$ values. Subscript "c" indicates age-constant hazards. The asymptotic value of the curve is λ/μ or 1, whichever is smaller. Apparent prevalence $\nu_c(t)$ equals the cumulative incidence $\pi(x)$ when the disease is benign ($\mu = 0$).

There is always some debate with respect to what it means to say that an animal is infected. In reality, the true infection state of an animal is usually difficult to define, let alone measure. We adopt a purely operational approach based on some prescribed diagnostic test outcome. We assume that a well-defined test is applied to the subject animal: the outcome of the test is either positive or negative. This test could be based on clinical signs, serology, immunohistochemistry, polymerase chain reaction (PCR), or some other methodology. It is assumed that the test outcome is associated (i.e., both sensitive and specific) with some meaningful underlying irreversible disease state, reasonably referred to as the state of "being infected." The extent to which the applied test is a good surrogate for this state of being infected will determine the biological interpretability of the results, although the analyses are always valid if the interpretation is restricted to just the narrow sense of test outcomes.

For ease of presentation, unless we state otherwise, it should be assumed that we are dealing with endemic disease situations, that is, situations in which the disease is essentially at equilibrium in the population or at least changing slowly relative to the life span of the host individual. Alternatively, if the data record is long enough, stable cycles about a mean can be accommodated as well (e.g., Farrington et al. 2001). This assumption is for ease of presentation only, and, as we briefly consider near the end of the *Discussion*, our approach is easily and naturally extended to non-equilbrium situations.

We start by introducing the notion of cumulative incidence. Although it would seldom be practical, one approach to measuring a disease's foothold in a population would be to introduce a cohort of negative sentinel animals into the population and then monitor them continuously to determine how rapidly they become positive (McCallum 2000). [For simplicity at this point, we will assume that any sentinel that dies before becoming positive would be replaced immediately with a new negative sentinel to carry on the monitoring of the original sentinel. The same result would be achieved by treating pre-conversion deaths as right-censoring events in an event-time analysis approach (Kalbfleisch and Prentice 1980), but the replacement assumption requires less explanation at this point.] For now, we assume that there is no age dependency with respect to the susceptibility to infection and that the population is homogenously mixed with respect to infectious contacts. If x is the exposure time of the sentinels in the population and disease conditions remain relatively stable, the expected proportion of the initial sentinel population that is still negative at time x after introduction will have decayed away exponentially according to $\exp(-\lambda x)$, where λ is the rate constant known as the force of infection, or infection hazard. The expected proportion of the sentinel cohort that is positive, the cumulative incidence, at time x is $\pi(x) = 1 - \exp(-\lambda x)$ (Fig. 1). The clearest way to think of the force of infection (λ) is that over a short interval of time from x to $x + \Delta$, the probability that a negative animal becomes positive is $\lambda\Delta$. (Initially, the cumulative incidence curve is well approximated with the linear relationship $\pi(x) \approx \lambda x$. A constant fraction of the remaining susceptible sentinels is constantly being converted, and the nonlinearity occurs as the pool of remaining susceptible sentinels gets progressively diminished; Fig. 1). The defining feature of cumulative incidence is the (almost always impractical) continuous longitudinal monitoring of all of the introduced sentinels.

If all animals are born test-negative, births have the effect of introducing negative sentinel cohorts into the natural population. This should immediately point out the source of age dependency in overall apparent prevalence. The cross-sectional sample on which apparent prevalence is based can be thought of as a mixture of sentinel cohort populations with various exposure times (ages); hence, the apparent prevalence will reflect this mix of ages. Populations skewed toward younger age classes, say by high reproduction, will have experienced less exposure time and, hence, will have a lower overall apparent prevalence.

Although cumulative incidence is a useful device for introducing the notion of the force of infection and illustrating why apparent prevalence shows strong age dependency, the cumulative incidence relationship $\pi(t) = 1 - \exp(-\lambda t)$ typically will not be especially useful in wildlife applications that result in age-specific apparent

prevalence data, where t now indicates the sampled subject's age. The reason for this is that the cross-sectional sampling on which apparent prevalence is based does not obtain a representative sample of all animals that were originally introduced into the sentinel birth cohorts because of the biasing effects of infection-associated mortality.

Infection-associated mortality is defined as the increase in mortality hazard resulting from becoming test-positive relative to an animal that remains test-negative. (It does not imply death due to any particular cause, and should not be confused with cause-specific mortality.) Infection-associated mortality has no effect on longitudinally sampled cumulative incidence; a sentinel animal is followed until it becomes positive, and whatever happens to the animal thereafter is irrelevant. In contrast, infection-associated mortality can exert a huge effect on cross-sectionally sampled apparent prevalence. Infection-associated mortality preferentially removes previously converted animals before a cross-sectional sample is obtained and, hence, the animals available for sampling will be biased toward negative animals. Thus, unlike cumulative incidence, a meaningful analysis of apparent prevalence and estimation of the force of infection must simultaneously accommodate infection-associated mortality. In general event-time analysis, an event, such as death, that prevents a subject from subsequently being recruited into the study is referred to as left truncation (Cox and Oakes 1984). In the wildlife and ecological literature, it is not uncommon for left censoring and left truncation to be confused. In left censoring, the subject is recruited into the study but the endpoint, such as infection, has already occurred by the time of study entry. In left truncation, the previous occurrence of some endpoint, usually death, prevents us from even being aware of the subject. Apparent prevalence data are both left-truncated by death and left-censored with respect to infection time. The effect of infection-associated mortality on left truncation substantially complicates the analysis of apparent prevalence data and prevents it from being an otherwise standard left-censored data problem in which left truncation can be ignored (e.g., Farrington et al. 2001).

As humans continue to encroach on natural ecosystems, the trend for infectious wildlife diseases to involve controversial human or domestic livestock aspects will increasingly require the attention of resource managers and researchers. Our goal is to acquaint readers with the central concepts and tools of quantitative epidemiology of infectious wildlife diseases, in which the force of infection plays a central role. This paper focuses on developing and comparing general force-of-infection models for apparent prevalence data for nonbenign diseases (i.e., diseases associated with an increase in death rate) which heretofore have been an essentially intractable problem. Following the standard approach in modern medical biostatistics, we formulate the problem in terms of hazard functions (e.g., Anderson et al. 1992) from which we then obtain maximum likelihood solutions. This provides a natural proportional-hazards framework for the regression analysis of variables of interest, which we describe and compare with better known logistic regression procedures. The age-prevalence approach to estimating the force of infection has been used with substantial success in human populations for many years (e.g., Anderson and May 1985, Grenfell and Anderson 1985, Bundy et al. 1987, Keiding 1991, Keiding et al. 1996, Farrington et al. 2001), including an until recently somewhat overlooked general treatment by Daniel Bernoulli in 1766 (Dietz and Heesterbeek 2002). Presence–absence data such as age-prevalance data are generally referred to as current status data, and the analysis of general current status data has received a substantial amount of attention (e.g., Jewell and van der Laan 2004). However, the vast majority of these analyses assume that any infection-associated mortality is negligible. There have been relatively few age-specific prevalence applications in zoology and wildlife; notable examples include transmission of nematodes in Red Grouse (Hudson and Dobson 1997), schistosome infections in snails (Cohen 1973, Woolhouse and Chandiwana 1992), and *Mycobacterium bovis* infection of ferrets (Caley and Hone 2002). Without exception, when any of these approaches allow for infection-associated mortality, only very simplistic force-of-infection models could be implemented, which we address. The methods are equally applicable, regardless of the infectious agent: prion, microorganism, or macroparasite.

Logistic Regression, Link Functions, and the Three-State Irreversible Disease Model

Apparent prevalence is an example of binary data. Each animal is tested for disease, and the outcome is that the animal is either positive ($Y = 1$) or negative ($Y = 0$). Many ecologists are familiar with logistic regression, which is a popular method for analyzing binary data (e.g., Trexler and Travis 1993), and indeed logistic regression has been the method of choice for analyzing wildlife apparent prevalence data. Logistic regression is a specific example of binary regression, where factors of interest (covariates), say the vector $\mathbf{X}$, are related to the probability of the event of interest, $\Pr(Y = 1 \mid \mathbf{X})$, through some function, the link function. (For the sake of technical correctness, we point out that what we subsequently will refer to as link functions are actually inverse link functions in the usual generalized linear model terminology.) In logistic regression, the link function is the logistic function:

$$\Pr(Y = 1|\mathbf{X}) = \frac{\exp(\beta_0 + \mathbf{X}\boldsymbol{\beta})}{1 + \exp(\beta_0 + \mathbf{X}\boldsymbol{\beta})}.$$

In the majority of applications of logistic regression, and wildlife disease prevalence is such an example, there is nothing especially natural about the logistic function. It is just a convenient way to make sure that predicted

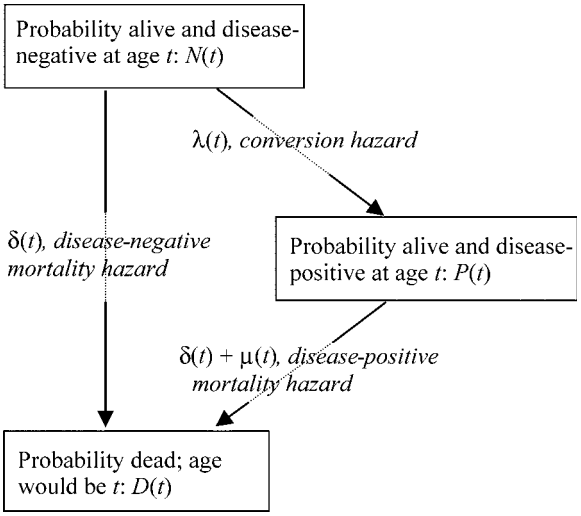

FIG. 2. The three-state irreversible disease model.

values for $\Pr(Y = 1 \mid \mathbf{X})$ stay within the feasible range of 0–1, and much of its popularity simply reflects the wide availability of software. The log odds ratio parameters, $\boldsymbol{\beta}$, have no direct dynamic epidemiological interpretation, and as will be described, the logistic function itself does not behave as apparent prevalence data are observed to behave. Our intention is not to criticize logistic regression, but rather to use it as a familiar starting point and to introduce the general notion of a link function, showing why alternatives are desirable. Our next step is to illustrate how a consideration of the disease process leads to alternative link functions that allow us to proceed in a manner quite similar to usual logistic regression estimation and modeling, with the important difference that the parameters and models correspond to specifically defined epidemiological hypotheses.

Fig. 2 depicts the three-state irreversible disease model (e.g., Keiding 1991, Anderson et al. 1992). Wildlife diseases such as chronic wasting disease, tuberculosis, brucellosis, and West Nile virus arguably fit this model. An animal's probability of being alive and disease-negative at age t is $N(t)$; of being alive and disease-positive at age t is $P(t)$; and of being dead at age t is $D(t)$. It is assumed that animals are born disease-negative. That is, at age $t = 0$, a birth cohort appears in the top compartment (alive and disease-negative) so $N(0) = 1$, $P(0) = 0$, and $D(0) = 0$. However, some animals may immediately start to leave this state by becoming positive and moving to the second state, P, or dying directly, moving to the third state, D. The age-specific hazard functions $\lambda(t)$, $\delta(t)$, and $\delta(t) + \mu(t)$ control the rates (probabilistically speaking) at which animals make transitions from one state to another. The force of infection, or infection hazard, $\lambda(t)$, is just as described before, except that we now allow for the possibility that it may vary with age t. The probability per unit time at which living, negative animals die is controlled by the infection-negative force of mortality, or mortality hazard, $\delta(t)$. The infection-associated mortality hazard is $\delta(t) + \mu(t)$, and in particular $\mu(t)$ reflects how infection increases the mortality hazard above the disease-negative rate $\delta(t)$.

Apparent prevalence is obtained from a cross-sectional sample of living animals; that is, samples are drawn from only the top two compartments, N and P of Fig. 2. Age-specific apparent prevalence is the probability that an animal, conditional on being alive at age t, is disease-positive, and is given as

$$\nu(t) = \frac{P(t)}{P(t) + N(t)}$$

which is the probability of being in the second compartment (P), conditional on being in one of the top two compartments. This specifies a link function $\Pr(Y = 1 \mid t) = \nu(t)$, which links the probability of the observation Y to an underlying model formulated in terms of hazard functions. The hazard functions in Fig. 2 essentially specify a system of differential equations that can be solved for the state probabilities $P(t)$, $N(t)$, and, in turn, $\nu(t)$ (see Appendix). In particular, if we assume that $\lambda(t)$ and $\mu(t)$ do not vary with age, that is, $\lambda(t) = \lambda$ and $\mu(t) = \mu$, we obtain the link function for age-specific apparent prevalence as

$$\nu_c(t) = \frac{1 - \exp[-(\lambda - \mu)t]}{1 - \frac{\mu}{\lambda}\exp[-(\lambda - \mu)t]} \qquad (1)$$

where the subscript "c" of $\nu_c(t)$ reminds us that the hazards are age-constant. Eq. 1 gives the age-specific apparent prevalence, correcting for the potentially higher mortality rate of positive animals.

This particular model has a remarkable history; with minor modifications, this model was presented by Daniel Bernoulli in 1766 (Seal 1977, Dietz and Heesterbeek 2002). This is a special case of Cohen's (1973) model when the disease cure rate ("defection") is 0. The behavior of this model for varying values of μ is shown on Fig. 1. When the infection-associated mortality-hazard increment is greater than the force of infection ($\mu > \lambda$), apparent prevalence $\nu_c(t)$ will never reach 1, but levels out at λ/μ. Note that when there is no infected-associated mortality ($\mu = 0$), the apparent prevalence model and the cumulative incidence model are equivalent: $\nu_{c0}(t) = 1 - \exp(-\lambda t)$. It is interesting that the infection-negative mortality hazard $\delta(t)$ has no direct bearing on these models (see Appendix), which follows from Cohen's (1972) observation. We now describe how link functions such as $\nu_c(t)$ can be used for analyzing apparent prevalence, and how variables of interest can be modeled in this context.

PROPORTIONAL INFECTION HAZARD MODELS

Because hazard functions are nonnegative functions, it is natural to model them with log-linear models.

Consider a vector of interesting covariates, say $\mathbf{X}$. Let $h_0(t)$ be the value that the hazard function assumes when all of the covariates are null ($\mathbf{X} = \mathbf{0}$); $h_0(t)$ is the baseline hazard function. A natural way to model the influence of covariates on a hazard function is then $\ln h(t) = \ln h_0(t) + \mathbf{X}\boldsymbol{\beta}$, or, equivalently, $h(t) = h_0(t)\exp(\mathbf{X}\boldsymbol{\beta})$. This is the basis for the proportional hazards model, popularized by Cox (1972). The term $\exp(X_i\beta_i)$ for the ith covariate X_i is referred to as a hazard ratio and the unit hazard ratio $\exp(\beta_i)$ measures how much a unit change of X_i shifts the hazard $h(t)$ up [$\exp(\beta_i) > 1$] or down [$\exp(\beta_i) < 1$].

Such models are easy to build into the structure of the Bernoulli/Cohen apparent prevalence link function $\nu_c(t)$. Any regression-type model $\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$ that would be appealing to examine, say with logistic regression, is just as readily accommodated by $\nu_c(t)$ by substituting $\lambda = \exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n)$ into $\nu_c(t)$. The resulting parameter estimates have specific epidemiological interpretations. The intercept β_0 is the log baseline force of infection, $\ln \lambda_0$, and the coefficients, β_i, are the log infection hazard ratios. In general, it is not appropriate to treat age t as just another covariate X_j. Note that the link function $\nu_c(t)$ already includes age t; if we include age t as a covariate X_j in the log-linear model for λ as well, we are essentially saying that our original assumptions about the age structure of $\lambda(t)$ used when we solved Fig. 2 for $\nu_c(t)$ need some correction. Such a model with t entered as a covariate X_j will not correspond to a known solution of Fig. 2, and hence should not be given a hazards-based interpretation. However, such models are very useful diagnostically; if age t (or some transform) as an ordinary covariate improves model fit substantially, it indicates that the age dependency of the baseline infection hazard is incorrectly specified. This then requires an exploration of models for $\nu(t)$ other than $\nu_c(t)$ if we want to preserve the hazards interpretation, which is generally a desirable goal. We consider this next.

Extension To Age-Varying Infection Hazard Models

In many circumstances, age invariance of the infection hazard required for $\nu_c(t)$ is unrealistic. The assumption that the infection hazard is age invariant results in the infection times following an exponential distribution. For general event-time analysis, numerous age-varying hazard functions have been proposed, and these make a reasonable starting point for age-varying infection hazard models. Popular two-parameter age-varying hazard models include the Weibull $\lambda(t) = \alpha\beta(\alpha t)^{\beta-1}$, the Gompertz $\lambda(t) = e^{\alpha+\beta t}$, the Pareto $\lambda(t) = \alpha\beta/(1 + \beta t)$, and the log-logistic $\lambda(t) = \alpha\beta(\alpha t)^{\beta-1}/[1 + (\alpha t)^{\beta}]$, where α and β are parameters that control the age dependency. These models arise from various assumptions about the event process (e.g., Hougaard 2000). The Weibull model, which contains the exponential model as a special case ($\beta = 1$), models the hazard as a power function of age. The Gompertz model also includes the exponential as a special case ($\beta = 0$) and assumes that the log hazard is linear with respect to age. The Pareto model is an example of a frailty, or heterogeneous hazard, model; the infection hazard is assumed to be age invariant for each animal i, $\lambda_i(t) = \lambda_i$, and λ_i for animal i is assumed to be the realization of a gamma random variable. Like the Gompertz and Weibull, the Pareto hazard is monotonic with respect to age. The log-logistic hazard is unique among these models in that it is not monotonic and can have a peak.

Unlike for $\nu_c(t)$, it appears to be impossible to find closed-form solutions for the age-specific apparent prevalence $\nu(t)$ for these hazard models (e.g., Caley and Hone 2002). However, we have developed a robust approximation method that allows one to compute $\nu(t)$ for any age-varying hazard function. The approach is based on the piecewise constant hazards (PCH) model $\lambda_{pc}(t)$, which divides the age axis into intervals $0 = t_{(0)} < t_{(1)} < \cdots < t_{(j)}$. During the qth interval from $t_{(q-1)}$ to $t_{(q)}$, the hazard remains constant at λ_q, so $\lambda_{pc}(t) = \sum_{q=1}^{m} \lambda_q 1\{t_{q-1} < t \leq t_{(q)}\}$where $1\{\ \}$ is the 0–1 indicator function (Hougaard 2000). The PCH is popular in event-time analysis because it is a flexible model that can approximate many different forms. In the Appendix, we give the closed-form solution for $\nu(t)$ for the PCH infection hazard model for any number of intervals, $\nu_{pc}(t)$. In turn, the PCH model $\nu_{pc}(t)$ can be used to approximate $\nu(t)$ for any infection hazard model $\lambda(t)$ by setting $\lambda_q = \lambda((t_{(q-1)} + t_{(q)})/2)$, the value of $\lambda(t)$ at the midpoint of the interval. This approach is very robust and, with an adequate number of intervals, achieves a high degree of accuracy (Appendix). Proportional-hazards models are easily incorporated into this framework by modeling λ_q as $\lambda_q = \lambda_{0q}\exp(\mathbf{X}\boldsymbol{\beta})$, where λ_{0q} is the midpoint hazard $\lambda((t_{(q-1)} + t_{(q)})/2)$ for the particular parametric model of interest.

Model Diagnosis and Comparison

For model fitting, we take a maximum likelihood approach, although Bayesian approaches could be used as well. For models fitted by maximum likelihood, there are two popular approaches for model comparison: information theoretic measures and likelihood ratio tests. If one model can be obtained from a more general model by imposing restrictions on the more general model, likelihood ratio statistics can be used to compare such models (McCullagh and Nelder 1983). Loosely speaking, the negative log maximum likelihood, $-L$ (see Appendix) can be viewed as a measure of the distance between the observed data and the fitted values. Let L_g and L_r be the log maximum likelihoods for the general and restricted models, respectively. For nested models it is always the case that $L_g > L_r$, even if the general model contains needless parameters, so the issue becomes whether the difference $L_g - L_r$ is large enough to suggest that the general model is a real improvement over the restricted model. For moderate sample sizes, under the null hypothesis that the additional parameters in the general model are unnecessary, the statistic $G^2 = 2(L_g -$

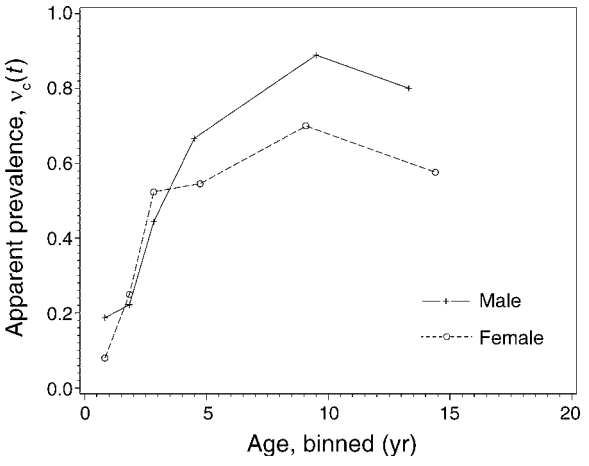

FIG. 3. Age-specific apparent prevalence by sex in bison (*Bison bison*). Bins were created for ages 3–5, 6–10, and 11 years and greater. The average age within the bin was used as the plotting point on the *x*-axis.

L_r) will follow a chi-squared distribution with "df" degrees of freedom, where "df" is the difference in the number of parameters between the two models. If a "significant" difference is not seen to exist between the two models, this is taken as evidence that the additional complexity of the general model is not warranted unless ancillary information suggests otherwise. Likelihood ratio tests are appropriate only for nested models, but information-theoretic approaches can be used to compare nested and nonnested models (Burnham and Anderson 1998). Akaike's Information Criteria (AIC) is defined as $\text{AIC} = 2(-L + q)$, where q is the number of parameters in a model and essentially penalizes the fit measured by $-L$ for the complexity of the model. Among competing models, the model with the smallest AIC is considered the "best." The differences between the AICs for competing models, ΔAIC, can be used to judge the relative strengths of support. Models with $\Delta\text{AIC} < 2$ relative to the best model need to be considered as serious contenders for best model, whereas models with $\Delta\text{AIC} > 4$ deserve substantially less credibility, and $\Delta\text{AIC} > 10$ indicates virtually no support (Burnham and Anderson 1998). However, as we note in our example, it is important that this approach be applied only to the set of those models that initially enjoy relatively comparable scientific support.

In ordinary regression, residuals are an important tool for diagnosing model adequacy. The extreme discreteness of binary data makes residual analysis problematic unless data can be grouped, which is indeed often feasible in wildlife apparent prevalence data. A "group" is composed of all animals that have the same age and covariate profile, such as sex and location. If n_k is the total number of subjects in group k, and y_k is the number that are positive, then the deviance residual (McCullagh and Nelder 1983) for group k is defined as follows:

$$d_k = \text{sign}(y_k - \hat{y}_k)\left[2y_k\log\left(\frac{y_k}{\hat{y}_k}\right) + 2(n_k - y_k)\log\left(\frac{n_k - y_k}{n_k - \hat{y}_k}\right)\right]^{1/2}.$$

Plotting d_k against age t or against covariates in $\mathbf{X}$ is useful for assessing ill-fitting model components, which we illustrate in our example. Deviance residuals essentially decompose the log likelihood L, and viewing the log likelihood as an overall-fit distance measure, the deviance residuals represent the individual group contributions to the fit.

For models such as we present, the distributions of the maximum likelihood estimates (MLEs) of the parameters will converge to normal distributions as the sample size grows. Thus, for example, approximate 95% confidence intervals can be constructed in the usual way about any parameter estimate as $\hat{\theta} \pm 1.96$ SE $(\hat{\theta})$. However, if the distribution of $\hat{\theta}$ is skewed, as might be expected for parameters restricted to be nonnegative, such as μ, convergence to normality may be slow. This can be improved by applying monotonic transformations thought to make the distribution more symmetric, such as $\ln(\hat{\theta})$ (e.g., Therneau and Grambsch 2000). A profile likelihood approach to confidence intervals, which is usually more exact and invariant to monotonic parameter transformation, is preferable (e.g., Therneau and Grambsch 2000). This approach is based on the likelihood ratio statistic $G^2 = 2(L_g - L_r)$, where the general model treats θ as a free parameter to be estimated and as a known constant in the restricted model. Then the 95% profile likelihood confidence set for θ are all of those values of θ for which $3.84 > 2(L_g - L_r)$; we illustrate this approach in the example.

EXAMPLE

Joly and Messier (2004) examined factors affecting the apparent prevalence of bovine tuberculosis (*Mycobacterium bovis*, BTB) in bison (*Bison bison*) in Wood Buffalo National Park, Canada; we use a subset of their data for illustration. We pool the Delta and Hay Camp bison and exclude the Nyarling River animals because no males were tested there. Our subset includes a total of 322 animals, of which 156 tested positive for TB (see Joly and Messier [2004] for details on capture and testing methods and diagnostic test criteria). The oldest animal was over 20 years old, and to generate reasonably smooth apparent prevalence plots, we bin some ages (Fig. 3).

The three primary questions of interest are (1) is the force of infection dependent on age, (2) is the force of infection sex specific, and (3) is there evidence for infection-associated mortality? For every "baseline" force-of-infection model, we fit four model structures. The "null" model assumes that there is no infection-associated mortality ($\mu = 0$), and furthermore assumes that there is no sex effect. The "sex" model assumes that there is no infection-associated mortality ($\mu = 0$), but that sex acts in a proportional hazards fashion on the

TABLE 1. Summary of model fits as measured by incremental AIC and likelihood ratio tests.

Model no.	Infection hazard	ΔAIC			
		Null	Sex	μ	Sex, μ
1	constant	36.7	32.7	7.8	7.5 (0.133, <0.001)
2	"constant," linear age	2.8	3.3	4.7	5.2 (0.216, 0.754)
3	"constant," quadratic age	−2.0	−1.8	−1.7	−1.2 (0.199, 0.204)
4	constant, year specific	40.5	36.5	11.2	11.1 (0.149, <0.001)
5	Weibull	14.6	15.1	9.8	9.5 (0.134, 0.006)
6	Pareto	10.1	10.2	6.5	6.6 (0.169, 0.018)
7	Gompertz	5.9	6.2	5.0	5.2 (0.187, 0.088)
8	log-logistic	11.7	10.9	0.7	0.0 (0.099, <0.001)

Notes: Four model structures were fitted for each infection hazard function assumption: "null" models assume that the infection hazard is not sex specific, and that conversion does not influence mortality (i.e., $\mu = 0$); "sex" models assume that the infection hazard is sex specific, but that conversion does not influence mortality ($\mu = 0$); "μ" models assume that the infection hazard is not sex specific, but that conversion does influence mortality; "sex, μ" models assume that the infection hazard is sex specific and that conversion does influence mortality. The log-logistic "sex, μ" model was the model to which all other 31 models were compared to obtain the ΔAIC values; its AIC was 389.97. The ΔAIC values are useful for ranking all 32 models: the smaller the ΔAIC value, the better the model. Within an infection hazard model, the parenthetical values in the "sex, μ" column are the *P* values resulting from likelihood ratio tests measuring whether sex improves a model already adjusted for μ, and whether μ improves a model already adjusted for sex, respectively. For example, adding sex to the log-logistic model that already includes μ results in a *P* value of 0.099 for sex. Models 2, 3, and, to some extent, 4 are diagnostic models and do not correspond to known hazards-based models; μ lacks a hazard interpretation. All of these models were fitted with the R function in the Supplement; models 1–4 were also fitted with SAS PROC NLMIXED and obtained very similar results.

infection hazard. The "μ" model assumes that there is infection-associated mortality, but no sex effect. The "sex, μ" model assumes that there is both infection-associated mortality and a sex effect on the infection hazard. As shown in Table 1, we fit eight baseline force-of-infection models; as will be discussed, some of these are "diagnostic" models and some are "scientific" or "hypothesis" models. Model comparisons are performed with likelihood ratio tests and the ΔAIC statistic.

Models assuming an age-constant infection hazard, $\nu_c(t)$, are especially tractable because they have closed-form likelihoods (Eq. 1) and so can be fitted with standard generalized nonlinear modeling software, such as SAS PROC NLMIXED (SAS Institute, Cary, North Carolina, USA). The deviance residuals for the age-constant (model 1) "μ, sex" model suggests a quadratic lack of fit with respect to age (Fig. 4). To further explore this, we fit

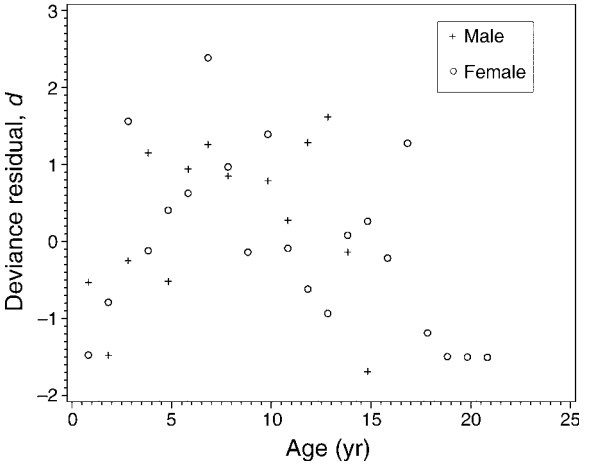

FIG. 4. Deviance residuals resulting from the proportional infection hazards Bernoulli/Cohen model $\nu_c(t)$ with a sex effect.

$\nu_c(t)$ models of the form $\ln \lambda = \beta_0 + \beta_1 t$ (model 2; linear age) and $\ln \lambda = \beta_0 + \beta_1 t^2$ (model 3; quadratic age). We used the best-fitting hazard-based model as the reference model for computing the ΔAICs, which is the log-logistic "μ, sex" model for which ΔAIC becomes 0. The ΔAICs allow us to rank all 32 models, and in this respect the quadratic models (model 3) fit somewhat better than the best hazard-based model, log-logistic "μ, sex." The good fit of models such as 2 or 3 present a dilemma; as noted previously, including age as a covariate causes the model to lose a hazards-based interpretation. Although models such as 2 and 3 are useful "curve-fitting" models in much the same spirit as logistic regression, they no longer represent known solutions to the three-state disease model. Depending on the goals, such as evaluating the relative importance of covariates, analyses based on models such as 2 or 3 might be adequate. However, developing an age-varying infection hazard model that preserves the proportional hazards interpretation will generally be desirable. Similar to models 2 and 3, model 4 is also a diagnostic model that was intended to examine whether prevalence patterns differed between the sampling years, e.g., due to sampling issues. The model fits give no evidence for between-year variation.

In general, proportional infection hazard models with age-varying infection hazards will not have closed-form likelihoods; to fit such models, we wrote an R function (R Development Core Team 2005) given in the Supplement, based on our piecewise constant hazard (PCH) approximation method (Appendix). The ΔAICs allow us to rank all 32 models. In general, no members of the Weibull (model 5), Pareto (model 6), or Gompertz (model 7) families of models result in especially good fits, but the log-logistic models (8) including μ fit nearly as well as the quadratic diagnostic models (3). The residuals from the log-logistic μ,sex model indicate an

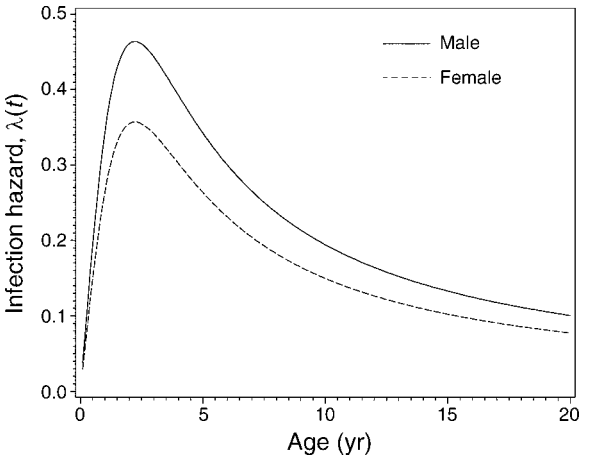

FIG. 5. Estimated force of infection (infection hazard) for the log-logistic model.

improved fit (not shown), no doubt because of the modal nature of the log-logistic hazard function (Fig. 5).

Within an infection hazard model family, likelihood ratio tests are useful for judging the relative contributions of μ or sex. From Table 1 it can be seen that including μ (infection-associated mortality hazard) in the valid hazard models generally improves the fit within a model family, and this is the case, in particular, for the log-logistic model (likelihood ratio $P < 0.001$). The perhaps most useful interpretation of μ is when it is transformed as exp(−μ), which gives the annual survival ratio for a positive animal relative to a negative animal. That is, for a subject alive and negative at r, its probability of surviving from r to $r+1$, given that it remains negative, is $S_\delta(r, r+1)$. Similarly, the survival to $r+1$ for a subject alive and positive at r can be shown to be $\exp(-\mu)S_\delta(r, r+1)$ (see Appendix); hence, the interpretation of exp(−μ) as the annual survival ratio. Care must be taken not to overinterpret μ, and, in this respect, it is useful to consider it within the context of competing risk theory (Kalbfleisch and Prentice 1980). The mortality hazard function for a negative animal is $\delta(t)$ and the mortality hazard rate for a positive animal is $\delta(t) + \mu$. The hazard increment μ represents the disease-associated mortality hazard in the context where all other sources of mortality are operating as well. It is usually untenable to conjecture that μ is independent of the general mortality hazard background; for example, predators may preferentially take positive animals, so reducing general predation intensity might have the effect of reducing μ as well. Although it may seem unsatisfying, generalizing μ to situations beyond those in which it was measured is conjectural. This is analogous to the problem of "risk removal" in the analysis of cause-specific mortality (Kalbfleisch and Prentice 1980, Heisey and Fuller 1985).

Fig. 6 shows the profile likelihood for the annual survival ratio exp(−μ) for the log-logistic model. The maximum likelihood estimate of exp(−μ) is 0.77, with a 95% profile likelihood confidence interval of 0.67–0.89. The estimate of exp(−μ) is rather sensitive to the proposed infection hazard model under which it is estimated. For the valid infection hazard models (1, 5, 6, 7, 8), estimates of exp(−μ) range from 0.64 to 0.86. The model-averaged estimate of exp(−μ) is 0.77 when averaged over all valid infection hazard models that included μ (Buckland et al. 1997).

Although the evidence for age dependence of the infection hazard seems clear, the evidence for sex specificity is equivocal (likelihood ratio $P = 0.099$). None of the model fits improved dramatically when a sex term was included (Table 1), although it did not seem to detract from the models either. In the log-logistic model, the infection hazards ratio for females relative to males was estimated to be 0.77, the effect of which is displayed on Fig. 5. The 95% profile likelihood confidence interval ranges from 0.57 to 1.05, including the null value of 1, consistent with the P value of 0.099.

DISCUSSION

Statistical model evaluation and selection must achieve a balance between prior belief or knowledge and strength of evidence conveyed by data. Clearly, models based on substantive underlying theory, with parameters interpretable within that theory, are preferable to atheoretical "curve-fitting" models that fit equally well. Indeed, theory-based models are to be preferred over curve-fitting models even if the curve-fitting models fit somewhat better. A substantially better fit by a curve-fitting model should direct the researcher to reconsider the specific structure of the theory-based

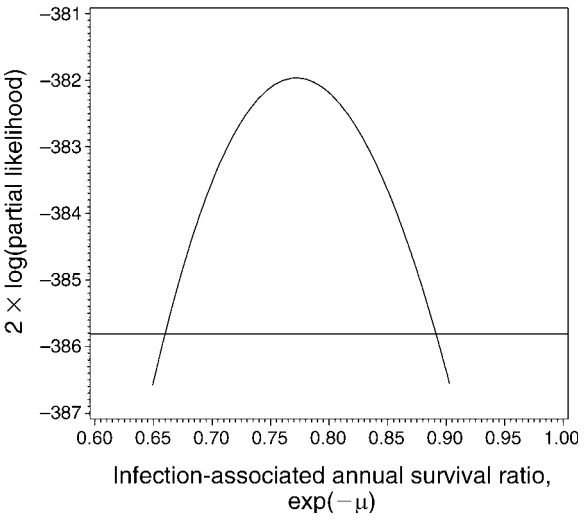

FIG. 6. Profile likelihood for the estimate of the infection-associated annual survival ratio exp(−μ). The partial likelihood was obtained by fixing μ [plotted on the x-axis as exp(−μ)] in the "μ, sex" log-logistic model 8 and then maximizing the likelihood for all of the remaining parameters. The y-axis is 2 × log(partial likelihood). The horizontal line is 3.84 units below the maximum likelihood and forms the cutoffs for a 95% CI. The best estimate (maximum likelihood) is 0.77, with a 95% CI of 0.67–0.89.

model, but seldom should the researcher prefer the curve-fitting model as a better representation of nature. In this respect, not all models are a priori equal, and model selection procedures such as Burnham and Anderson's (1998) AIC-based approach must be applied with enough flexibility to take this into account. Thus, although our "constant infection hazard, quadratic time" models produced the best fits from an AIC standpoint, we do not seriously entertain these as the "best," or even especially interesting, scientific models because they do not correspond with (known) solutions to the three-state model. (Of course our thinking on this might change if it were discovered that they were solutions under reasonable conditions.)

Caley and Hone (2002) show that simple piecewise force-of-infection models could be quite useful, but with their approach, they noted that obtaining solutions for even three-piece models involves "considerable complications." Our approach allows the easy fitting of piecewise models with any number of pieces, but we would generally recommend piecewise models only for identifying the general shape of other parametric models. AIC model selection for general piecewise models is problematic: of course each "piece" contributes a parameter for $\lambda(t)$, but if knot positions can be manipulated, as they usually can, they, too, should be viewed as parameters, thus complicating dimensionality considerations. A theoretically preferable approach might be to specify a piecewise model with many pieces and then impose a roughness penalty and use an information theoretical approach to measure the approximate dimensionality and select the appropriate degree of smoothing (e.g., Eilers and Marx 1996).

To our knowledge, this is the first application that demonstrates how the Bernoulli/Cohen model $\nu_c(t)$ and its extensions based on solutions of Fig. 1 can be used for proportional-hazards modeling of the force of infection. As noted, a special case of $\nu_c(t)$ occurs when the disease is assumed to be benign. This results in the model $\nu_{c0}(t) = 1 - \exp(-\lambda t)$, which has had a long history in modeling human disease prevalence data (e.g., Muench 1959, Draper et al. 1972, Griffiths 1974); Muench referred to it as the catalytic model. Today, this model is referred to as the interval-censored proportional-hazards model (Prentice and Gloeckler 1978) or the complementary log–log link model in generalized linear models. Such models are easily fitted with most logistic regression or generalized linear model packages (Allison 1997).

For the bison data, we have demonstrated strong support for the infection hazard of BTB being age dependent. At least two very different mechanisms can give rise to such patterns. Under a "homogeneous population" model, all animals have the same age-specific infection hazard patterns, which arise from age-specific physiological or behavioral mechanisms common to all animals. Under a "heterogeneous population," or frailty, model, each animal has a different infection hazard function. The event of interest occurs on average in the "most frail" high-hazard subjects first, with the result that the average hazard for the population eventually drops as the high-hazard individuals are eliminated. Thus, the force-of-infection pattern observed for bison may reflect one or some combination of these two mechanisms. An interesting result of this analysis is the demonstration of reduced survival in tuberculosis-positive bison. Joly and Messier (2005) were unable to demonstrate a main effect of disease on survival based on radiotelemetry data collected over a three-year period. The use of age-specific apparent prevalence data such as this may provide a more powerful test of the association of disease with survival rates in wildlife populations than can be obtained through radiotelemetry.

We have focused on the situation in which the infection hazard may depend on age, but not independently on calendar time. However, in epidemic rather than endemic situations, the role of calendar time may become very important in addition to age. Generalizing our approach to both calendar time and age presents no particular theoretical challenges (Keiding 1991), although estimation in the presence of both considerations will be more demanding data-wise. Let b be the calendar time of birth. Many models could be proposed for $\lambda(t, b + t)$, the bivariate age and calendar time-dependent infection hazard. A reasonable starting point is a separable model of the form $\lambda(t, b + t) = f(b + t)h(t)$ where $f(b + t)$ is some nonnegative function of calendar time and $h(t)$ is any age-dependent hazard function such as previous considered. With a minor change in notation, the derivation in the Appendix holds (Keiding 1991), and such models are very easily dealt with within our PCH approximation framework. However, unless data have been collected over some extended period of time, $b + t$ and t could be highly confounded, making it difficult or impossible to untangle age dependencies from calendar time dependencies. Joint calendar time/age modeling is thus most applicable to long-term, or panel (Jewel and van der Laan 2004), data sets. There are numerous circumstances in which wildlife disease data sets are georeferenced, and a natural direction for future research is to extend models such as we present here from the temporal domain to the spatiotemporal domain (Lawson 2001).

As the well-known statistician George Box is credited with saying, "All models are wrong, but some are useful." It should be recognized that our approach rests on several assumptions. A basic biological assumption is that the disease test is relatively good, i.e., the test has relatively good specificity and sensitivity through time, and that the notion of a conversion time is a meaningful idea for the particular disease and the particular test. Although the age-specific prevalence pattern suggests that the infection hazard is modal, other explanations for this pattern exist. Perhaps previously infected animals become less likely to test positive as they age; in bovids, individuals may become less likely to test positive as BTB progresses (Joly and Messier 2004).

Alternatively, the disease-associated mortality may be age specific, wherein the disease is more lethal in older animals. These issues are very difficult to untangle from just apparent prevalence data. In this respect, such analyses are valuable for guiding what questions should be examined in future studies and experiments.

Acknowledgments

Thanks to J. A. Ahumada, C. M. Bunck, P. W. Rasmussen, and M. D. Samuel for their critical reading and comments on earlier versions of the manuscript.

Literature Cited

Allison, P. D. 1997. Survival analysis using the SAS system. A practical guide. SAS Institute, Cary, North Carolina, USA.

Anderson, P. K., O. Borgan, R. D. Gill, and N. Keiding. 1992. Statistical models based on counting processes. Springer-Verlag, New York, New York, USA.

Anderson, R. M., and R. M. May. 1985. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. Journal of Hygiene **94**:365–436.

Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: an integral part of inference. Biometrics **53**: 603–618.

Bundy, D. A. P., E. S. Cooper, D. E. Thompson, R. M. Anderson, and J. M. Didier. 1987. Age-related prevalence and intensity of *Trichuris trichiura* infection in a St. Lucian community. Transactions of the Royal Society of Tropical Medicine and Hygiene **81**:85–94.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.

Caley, P., and J. Hone. 2002. Estimating the force-of-infection; *Mycobacterium bovis* infection in feral ferrets *Mustela furo* in New Zealand. Journal of Animal Ecology **71**:44–54.

Cohen, J. E. 1972. When does a leaky compartment model appear to have no leaks? Theoretical Population Biology **3**: 404–405.

Cohen, J. E. 1973. Selective host mortality in a catalytic model applied to schistosomiasis. American Naturalist **107**:199–212.

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

Cox, D. R., and D. Oakes. 1984. Analysis of survival data. Chapman and Hall, New York, New York, USA.

Dietz, K., and J. A. P. Heesterbeek. 2002. Daniel Bernoulli's epidemiological model revisited. Mathematical Biosciences **180**:1–21.

Draper, C. C., A. Voller, and R. G. Carpenter. 1972. The epidemiologic interpretation of serologic data in malaria. American Journal of Tropical Medicine and Hygiene **21**:696–703.

Eilers, P. H. C., and B. D. Marx. 1996. Flexible smoothing with B-splines and penalties. Statistical Science **11**:89–121.

Farrington, C. P., M. N. Kanaan, and N. J. Gay. 2001. Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. Applied Statistics **50**:251–292.

Grenfell, B. T., and R. M. Anderson. 1985. The estimation of age-related rates of infection from case notifications and serological data. Journal of Hygiene **95**:419–436.

Griffiths, D. 1974. A catalytic model of infection for measles. Applied Statistics **23**:330–339.

Heisey, D. M., and T. K. Fuller. 1985. Evaluation of survival and cause-specific mortality rates using telemetry data. Journal of Wildlife Management **49**:668–674.

Hougaard, P. 2000. Analysis of multivariate survival data. Springer-Verlag. New York, New York, USA.

Hudson, P. J., and A. P. Dobson. 1997. Transmission dynamics and host–parasite interactions of *Trichostrongylus tenuis* in red grouse *(Lagopus lagopus scoticus)*. Journal of Parasitology **83**:194–202.

Jewell, N. P., and M. J. van der Laan. 2004. Current status data: review, recent developments, and open problems. Pages 625–643 *in* N. Balakrishnan and C. R. Rao, editors. Handbook of statistics. Volume 23, Elsevier/North-Holland, Amsterdam, The Netherlands.

Joly, D. O., and F. Messier. 2004. Factors affecting apparent prevalence of tuberculosis and brucellosis in wood bison. Journal of Animal Ecology **73**:623–631.

Joly, D. O., and F. Messier. 2005. The effect of bovine tuberculosis and brucellosis on reproduction and survival of wood bison in Wood Buffalo National Park. Journal of Animal Ecology **74**:543–551.

Kalbfleisch, J. D., and R. L. Prentice. 1980. The statistical analysis of failure time data. John Wiley, New York, New York, USA.

Keiding, N. 1991. Age-specific incidence and prevalence: a statistical perspective. Journal of the Royal Statistical Society, Series B **154**:371–412.

Keiding, N., K. Begtrup, T. H. Scheike, and G. Hasibeder. 1996. Estimation from current-status data in continuous time. Lifetime Data Analysis **2**:119–129.

Lawson, A. B. 2001. Statistical methods in spatial epidemiology. John Wiley, New York, New York, USA.

McCallum, H. 2000. Population parameters. Blackwell Science, London, UK.

McCullagh, P., and J. A. Nelder. 1983. Generalized linear models. Chapman and Hall, New York, New York, USA.

Muench, H. 1959. Catalytic models in epidemiology. Harvard University Press, Cambridge, Massachusetts, USA.

Prentice, R. L., and L. A. Gloeckler. 1987. Regression analysis of grouped survival data with application to breast cancer data. Biometrics **34**:57–67.

R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Seal, H. L. 1977. Studies in the history of probability and statistics. XXXV. Multiple decrements or competing risks. Biometrika **64**:429–439.

Therneau, T. M., and P. M. Grambsch. 2000. Modeling survival data. Springer-Verlag, New York, New York USA.

Trexler, J. C., and J. Travis. 1993. Nontraditional regression analyses. Ecology **74**:1629–1637.

Woolhouse, M. E. J., and S. K. Chandiwana. 1992. A further model for temporal patterns in the epidemiology of schistomsome infections of snails. Parasitology **104**:443–449.

APPENDIX

State probability solutions for the three-state irreversible disease model (*Ecological Archives* E087-143-A1).

SUPPLEMENT

R function for fitting age-varying force-of-infection models to disease prevalence data (*Ecological Archives* E087-143-S1).