STATISTICS IN MEDICINE
Statist. Med. 2007; **26**:4352–4374
Published online 6 March 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.2836

TUTORIAL IN BIOSTATISTICS

Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution

Christopher Cox, Haitao Chu, Michael F. Schneider and Alvaro Muñoz*, †

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, U.S.A.

SUMMARY

The widely used Cox proportional hazards regression model for the analysis of censored survival data has limited utility when either hazard functions themselves are of primary interest, or when relative times instead of relative hazards are the relevant measures of association. Parametric regression models are an attractive option in situations such as this, although the choice of a particular model from the available families of distributions can be problematic. The generalized gamma (GG) distribution is an extensive family that contains nearly all of the most commonly used distributions, including the exponential, Weibull, log normal and gamma. More importantly, the GG family includes all four of the most common types of hazard function: monotonically increasing and decreasing, as well as bathtub and arc-shaped hazards. We present here a taxonomy of the hazard functions of the GG family, which includes various special distributions and allows depiction of effects of exposures on hazard functions. We applied the proposed taxonomy to study survival after a diagnosis of clinical AIDS during different eras of HIV therapy, where proportionality of hazard functions was clearly not fulfilled and flexibility in estimating hazards with very different shapes was needed. Comparisons of survival after AIDS in different eras of therapy are presented in terms of both relative times and relative hazards. Standard errors for these and other derived quantities are computed using the delta method and checked using the bootstrap. Description of standard statistical software (Stata, SAS and S-Plus) for the computations is included and available at http://statepi.jhsph.edu/software. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: survival analysis; left truncation; parametric models; generalized gamma distribution; non-proportional hazards; relative times

*Correspondence to: Alvaro Muñoz, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room E7648, Baltimore, MD 21205, U.S.A.
†E-mail: jvaldez@jhsph.edu

Contract/grant sponsor: National Institute of Allergy and Infectious Diseases; contract/grant numbers: UO1-AI-35042, UO1-AI-35043, UO1-AI-37984, UO1-AI-35039, UO1-AI-35040, UO1-AI-37613, UO1-AI-35041, UO1-AI-35004, UO1-AI-31834, UO1-AI-34994, UO1-AI-34989, UO1-AI-34993, UO1-AI-42590
Contract/grant sponsor: National Institute of Child Health and Human Development; contract/grant number: UO1-CH-32632
Contract/grant sponsor: National Center for Research Resources; contract/grant numbers: MO1-RR-00722, MO1-RR-00071, MO1-RR-00079, MO1-RR-00083

Received 19 January 2006
Accepted 21 December 2006

1. INTRODUCTION

A number of standard approaches are currently available for the analysis of censored survival data. In biomedical applications the semi-parametric Cox proportional hazards model is used extensively, and both the strengths and weaknesses of this approach are well known. Parametric models are attractive because standard methods such as maximum likelihood are available for parameter estimation and testing, the proportional hazards assumption is not required, and a complete description of the hazard function can be obtained [1]. A number of different parametric distributions are available, including the exponential, Weibull, log normal, gamma and log logistic. Use of parametric models facilitates the assessment of effects of exposures, not only by relative hazards, but also by relative quantiles (i.e. relative times: the ratio of times that a given percentage of individuals with different exposures take to develop the event of interest). The use of relative times has been hindered by the widespread use of semi-parametric proportional hazards models, which have intrinsic limitations for estimating and assessing the uncertainty of the relative times [2]. Relative times are an effective way to communicate inferences to the population at large, and thus have the potential for greater impact.

Each of the standard parametric distributions has hazard functions with different shapes, depending on the distribution and in most cases the values of the parameters as well. Choosing among the various alternatives can be difficult, yet the decision can have a considerable effect on the resulting inference. The generalized gamma (GG) [3, 4] is a parametric family that includes most of the commonly used distributions (e.g. exponential, Weibull, gamma, log normal). Prentice [5] proposed a parameterization for which maximum likelihood estimates could be computed using standard algorithms. As a result, this distribution is now available in standard statistical packages (e.g. Stata and SAS). Renewed interest in parametric models and the GG in particular is indicated by a recent article in a leading medical journal [6], and it seems timely to present a tutorial on the analysis of survival data using the GG family. In addition to reviewing the properties of the GG and illustrating its application using both relative times and relative hazards, the primary purpose of this tutorial is to enhance the usefulness of this broad family of distributions by providing a taxonomy for the behaviour of the hazard functions, with a graphical representation of its various members so that effects of covariates can be depicted. The taxonomy shows that the GG family includes all four of the common types of hazard function (i.e. increasing, decreasing, U and inverted-U shapes).

We have recently used Cox and Weibull regression models to describe the patterns of the hazard of death after clinical AIDS in different eras of antiretroviral therapy from 1984 to 2004 [7]. Here, we apply GG regression models to the data, unveiling the limitations of Cox and Weibull models, including lack of proportionality in both relative hazards and relative times. All computations were performed using standard statistical packages, as described and illustrated in Section 5. The complete data set and programs for the GG models discussed here are available at http://statepi.jhsph.edu/software

2. MODELS AND METHODS

2.1. The Generalized gamma distribution

2.1.1. Survival and density functions. The GG distribution is a three-parameter family with location (β), scale ($\sigma>0$) and shape (λ) parameters that generalizes the two-parameter gamma distribution. Specifically, let $\Gamma(t;\gamma)=\int_0^t x^{\gamma-1}\mathrm{e}^{-x}\,\mathrm{d}x/\Gamma(\gamma)$ be the cumulative distribution function (CDF) for

DOI: 10.1002/sim

the particular case of the gamma distribution with mean and variance equal to $\gamma>0$. The survival function for GG(β, σ, λ) is

$$S_{\mathrm{GG}}(t) = 1 - \Gamma[\lambda^{-2}\exp(\lambda[\log(t)-\beta]/\sigma);\lambda^{-2}] = 1 - \Gamma[\lambda^{-2}(\mathrm{e}^{-\beta}t)^{\lambda/\sigma};\lambda^{-2}] \quad \text{if } \lambda > 0 \tag{1}$$

and

$$S_{\mathrm{GG}}(t) = \Gamma[\lambda^{-2}(\mathrm{e}^{-\beta}t)^{\lambda/\sigma};\lambda^{-2}] \quad \text{if } \lambda<0 \tag{2}$$

From (1) and (2) the probability density function (PDF) is

$$f_{\mathrm{GG}}(t) = \frac{|\lambda|}{\sigma t\Gamma(\lambda^{-2})}[\lambda^{-2}(\mathrm{e}^{-\beta}t)^{\lambda/\sigma}]^{\lambda^{-2}}\exp[-\lambda^{-2}(\mathrm{e}^{-\beta}t)^{\lambda/\sigma}] \tag{3}$$

It follows from this definition that the special case $\lambda=\sigma$ is the two-parameter gamma distribution with mean equal to e^{β} and coefficient of variation equal to σ (i.e. variance $=\sigma^2\mathrm{e}^{2\beta}$). Hereafter, we denote the two-parameter gamma distribution by $G(\beta,\sigma)$ corresponding to the special case GG(β, σ, σ). The Weibull distribution corresponds to the case $\lambda=1$, and the very special case $\lambda=\sigma=1$ is the exponential distribution. The limiting case $\lambda=0$ is the log normal distribution, with $S_{\mathrm{GG}(\beta,\sigma,0)}(t) = 1-\Phi(\log(\mathrm{e}^{-\beta}t)^{1/\sigma})$, where $\Phi(\cdot)$ is the standard normal CDF. Thus, the GG family contains nearly all the most commonly used parametric distributions.

The GG family is closed under power transformations [4]. Namely, it follows from (1) and (2) that if $T\sim\mathrm{GG}(\beta,\sigma,\lambda)$ and $b\neq0$, then $\mathrm{e}^{a}T^{b}\sim\mathrm{GG}(a+\beta b,\sigma|b|,\lambda\,\mathrm{sign}(b))$. A special case is the inverse transformation, $1/T\sim\mathrm{GG}(-\beta,\sigma,-\lambda)$. In particular, the inverse Weibull ($\lambda=-1$), and inverse gamma ($\lambda=-\sigma$) are members of the GG family. If $T\sim\mathrm{GG}(0,1,\lambda)$ to which we refer as the standard ($\beta=0,\ \sigma=1$) GG for a given value of λ, then $\mathrm{e}^{\beta}T^{\sigma}\sim\mathrm{GG}(\beta,\sigma,\lambda)$. Furthermore, if $T\sim\mathrm{GG}(\beta,\sigma,\lambda)$, then $T^{\lambda/\sigma}\sim\mathrm{GG}(\beta\lambda/\sigma,|\lambda|,|\lambda|)$, corresponding to the gamma distribution $G(\beta\lambda/\sigma,|\lambda|)$. Hence, $S_{\mathrm{GG}(\beta,\sigma,\lambda)}(t)$ equals $S_{G(\beta\lambda/\sigma,\lambda)}(t^{\lambda/\sigma})$ for $\lambda>0$, and equals $1-S_{G(\beta\lambda/\sigma,-\lambda)}(t^{\lambda/\sigma})$ for $\lambda<0$. This relationship facilitates the application of the GG family since software that provides the cumulative distribution and density for the gamma distribution can be used to compute those of the GG, and therefore likelihood functions of data subject to several types of incompleteness (e.g. censoring, truncation) can be calculated and maximized (as discussed in Section 2.4). Implementation using standard statistical software is described in Section 5.

2.1.2. Quantile function. The time by which p per cent of the population experience the event (i.e. the pth percentile) is a useful statistic leading to a natural measure of association: relative times (Section 2.3). Let $\Gamma^{-1}(p;\gamma)$ denote the quantile function of the gamma distribution with mean and variance equal to $\gamma>0$; that is, $\Gamma(\Gamma^{-1}(p;\gamma);\gamma)=p$. Then from (1) and (2), the pth percentile of the standard ($\beta=0,\ \sigma=1$) GG, $t_{\mathrm{GG}(0,1,\lambda)}(p)$ is $[\lambda^2\Gamma^{-1}(p;\lambda^{-2})]^{1/\lambda}$ for $\lambda>0$, and $[\lambda^2\Gamma^{-1}(1-p;\lambda^{-2})]^{1/\lambda}$ for $\lambda<0$.

The quantile function for GG(β, σ, λ) satisfies

$$\log[t_{\mathrm{GG}(\beta,\sigma,\lambda)}(p)] = \beta+\sigma\log[t_{\mathrm{GG}(0,1,\lambda)}(p)] = \beta+\sigma g_{\lambda}(p) \tag{4}$$

where $g_{\lambda}(p)=\log[t_{\mathrm{GG}(0,1,\lambda)}(p)]$. It can be shown that $g_{\lambda}(p)=-g_{-\lambda}(1-p)$, which is consonant with the closure under inversion of the GG family. In particular, for the log normal ($\beta=0,\ \sigma=1,\ \lambda=0$), $g_0(p)=\Phi^{-1}(p)$, the quantile function of a standard normal variate; for the exponential ($\beta=0,\ \sigma=1,\ \lambda=1$), $g_1(p)=\log(-\log(1-p))$; and for the inverse exponential ($\beta=0,\ \sigma=1,\ \lambda=-1$), $g_{-1}(p)=-\log(-\log(p))$.

DOI: 10.1002/sim

Equation (4) characterizes the structure of the GG family, whereby λ determines the logs of the standard ($\beta=0,\ \sigma=1$) percentiles, which are then tilted by σ and shifted by β to provide the log percentiles of GG(β, σ, λ). Equation (4) also makes explicit the link between the GG and the conventional accelerated failure time (AFT) model [1], whereby if T_0 is a random variable whose distribution is GG(0, 1, λ), then $T=e^{\beta}T_0^{\sigma}$ (i.e. $\log(T)=\beta+\sigma\log(T_0)$) has distribution GG($\beta, \sigma, \lambda$). It should be noted that the classical AFT model holds only when covariate effects are modelled through the beta parameter. In this paper, we extend the analysis to covariates having effects through the sigma and/or lambda parameters and these are no longer AFT models.

With respect to links between the quantiles of the gamma and those of the GG, while $t_{G(\beta,\sigma)}(p)=e^{\beta}\sigma^2\Gamma^{-1}(p;\sigma^{-2})$ is the quantile function of the gamma distribution, (4) shows that $t_{\mathrm{GG}(\beta,\sigma,\lambda)}(p)$ is equal to $[t_{G(\beta\lambda/\sigma,\lambda)}(p)]^{\sigma/\lambda}$ for $\lambda>0$ and $[t_{G(\beta\lambda/\sigma,-\lambda)}(1-p)]^{\sigma/\lambda}$ for $\lambda<0$. Again, this facilitates the use of software which provides the quantile function for the gamma distribution to obtain the percentiles of the GG (see Section 5).

2.1.3. Interpretation of parameters. The location parameter β acts as a multiplier of time by the factor $e^{-\beta}$ in (3) and thus governs the value of the median for fixed values of σ and λ. Specifically, $\beta=\log(\text{median})-\sigma g_{\lambda}(0.5)$. The scale parameter σ determines the value of the interquartile ratio (i.e. IQr = 3rd quartile/1st quartile) for a fixed value of λ and independently of β. Specifically, $\sigma=\log(\mathrm{IQr})/\{\log[t_{\mathrm{GG}(0,1,\lambda)}(0.75)/t_{\mathrm{GG}(0,1,\lambda)}(0.25)]\}$. The parameter λ determines the percentiles of the standard ($\beta=0,\ \sigma=1$) GG. More importantly, as described in the next section, the parameters λ and σ together determine the *type* of hazard function and yield a graphical taxonomy (described here for the first time) of the rich variety of hazards covered by the GG.

2.2. Taxonomy of hazard functions

Since β simply acts as a multiplicative factor of time, we can present the hazard taxonomy for the GG family on the basis of the values of σ and λ. Indeed, if $h(t)$ is the hazard function for $T\sim\mathrm{GG}(\beta,\sigma,\lambda)$, then the hazard for $e^{-\beta}T\sim\mathrm{GG}(0,\sigma,\lambda)$ is $e^{\beta}h(e^{\beta}t)$. Thus, the location parameter accelerates the time and scales the overall level of the hazard, but does not influence the shape of the hazard function. The elements defining the taxonomy are the limiting values of the hazard as $t\to 0^+$ and $t\to\infty$, and the behaviour of the hazard between these two extremes which was considered by Glaser [8] for the case $\lambda>0$. Figure 1 shows that the two curves $\lambda=\sigma$ and $\lambda=1/\sigma$ divide the (σ,λ) half-plane into four regions within which hazard functions have the same type. For each region, Figure 1 shows the limiting values of the hazard functions in the form (a, b) where a and b are the limits of $h(t)$ as $t\to 0^+$ and $t\to\infty$, respectively. The arrows describe the behaviour (increasing with an upward arrow and decreasing with a downward arrow) of the hazard functions between these limits.

The hazard functions in the right-hand region of Figure 1 defined by $\{(\sigma,\lambda)|\sigma>1 \text{ and } (1/\sigma)\leqslant\lambda\leqslant\sigma\}$ are monotonically decreasing. In the interior of this region, they decrease from infinity to zero. On the boundary curve $\lambda=1/\sigma$, the hazard functions decrease from $c=\sigma^{2(\sigma^2-1)}e^{-\beta}/\Gamma(\sigma^2)$ to zero; and on the other boundary $\lambda=\sigma$, they decrease from infinity to $d=e^{-\beta}/\sigma^2$. In contrast, the hazard functions in the left-hand region of Figure 1, defined by $\{(\sigma,\lambda)|0<\sigma<1 \text{ and } \sigma\leqslant\lambda\leqslant(1/\sigma)\}$ are monotonically increasing from zero to infinity, except at the boundaries of the region, where they increase to d from zero if $\lambda=\sigma$, or from c to infinity if $\lambda=1/\sigma$. The hazard functions in the upper region of Figure 1 defined by $\{(\sigma,\lambda)|\lambda>\max\{\sigma,1/\sigma\}\}$ have a bathtub shape, unbounded in

DOI: 10.1002/sim

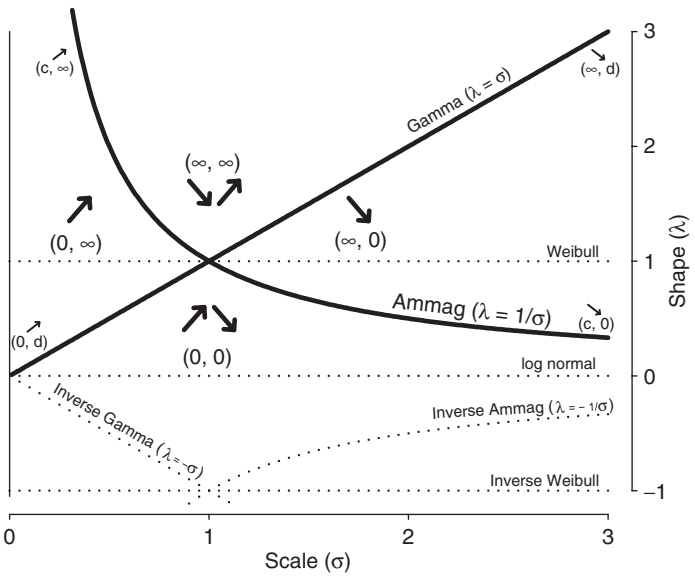


Figure 1. A schematic representation in the (σ, λ) half-plane of the generalized gamma distribution family. The four regions defined by the two curves include the four common types of hazard function: increasing and decreasing failure rate, bathtub and arc-shaped.

both directions of time. In contrast, the hazard functions in the lower region of Figure 1 defined by $\{(\sigma, \lambda)|\lambda < \min\{\sigma, 1/\sigma\}\}$ have an arc shape, starting and ending at zero.

The various special cases of the GG distribution can also be represented in the two-dimensional half-plane shown in Figure 1. The gamma and inverse gamma families are represented by the diagonal lines $\lambda = \sigma$ and $\lambda = -\sigma$, and the Weibull, log normal and inverse Weibull by the horizontal lines $\lambda = 1, 0$ and -1, respectively. The exponential distribution corresponds to the point $(\sigma, \lambda) = (1, 1)$. Another useful distribution is specified by the hyperbola $\lambda = 1/\sigma$, an essential element of the taxonomy. Because of the reciprocal relationship with the scale parameter of the gamma family, we will refer to this distribution as the *ammag*, which is *gamma* written backwards and its name and role are presented here for the first time. The inverse *ammag* is specified by $\lambda = -1/\sigma$.

Examples of the four types of hazard function, with $\beta = 0$ and values of (σ, λ) from each region, are shown in Figure 2. Decreasing and increasing hazard functions are illustrated in the two upper panels. The solid line in the upper left panel shows a decreasing Weibull hazard corresponding to $\sigma = \frac{4}{3} > 1 = \lambda > \frac{3}{4} = 1/\sigma$. The upper right panel illustrates increasing hazards; the solid line is also Weibull, for $\sigma = \frac{3}{4} < 1 = \lambda < \frac{4}{3} = 1/\sigma$. Bathtub and arc-shaped hazards are illustrated in the lower two panels of Figure 2; the solid line in the lower left panel has a bathtub shape corresponding to $\lambda = 4 > 1 = \sigma = 1/\sigma$, and the lower right panel has a solid line that is log normal, with $\sigma = \frac{3}{4}$. Figure 2 illustrates the rich diversity of available shapes for GG hazard functions.

2.2.1. The ammag distribution. The *ammag* distribution $A(\beta, \sigma)$ for $\sigma > 0$ corresponds to the particular case of $GG(\beta, \sigma, 1/\sigma)$, and is the only member of the GG family whose hazard function has a finite, positive left-hand limit ($t \to 0^+$). From equation (4), the logs of the percentiles of $A(\beta, \sigma)$

DOI: 10.1002/sim

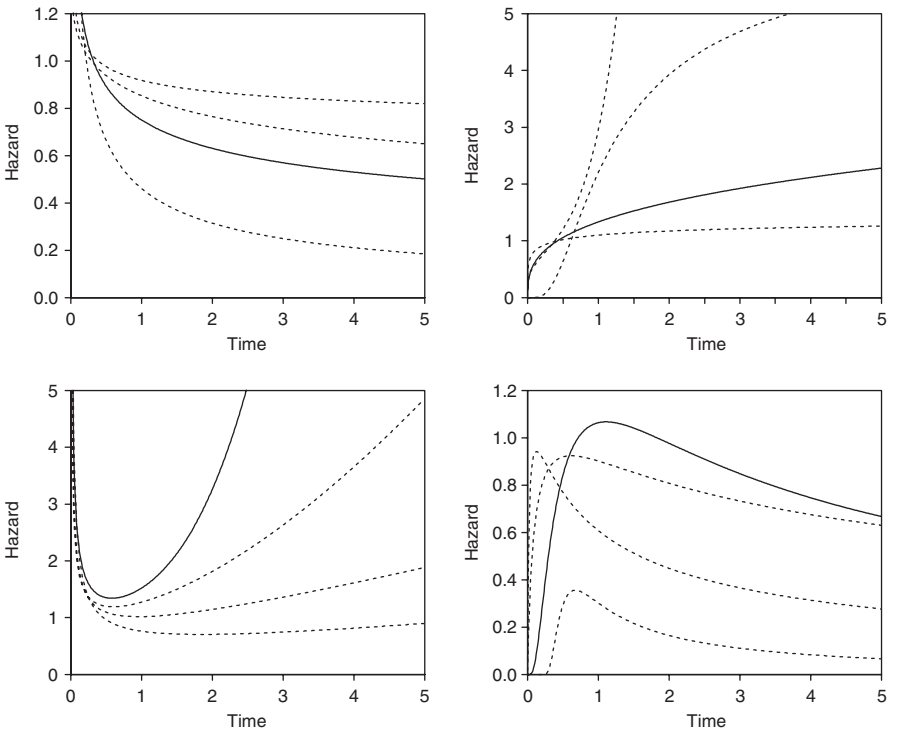


Figure 2. Examples of the four types of hazard function for the generalized gamma distribution, corresponding to the four regions in Figure 1.

can be written in terms of the percentiles of the gamma $G(0, 1/\sigma)$ as $\beta + \sigma^2 \log[t_{G(0,1/\sigma)}(p)]$. The availability of quantiles for the gamma distribution provides the elements to compute those for the *ammag*.

The *ammag* distribution is also instrumental for providing a taxonomy of the densities. Indeed, for $\lambda<1/\sigma$ (i.e. below the *ammag* in Figure 1), the density has a value of zero at $t=0$, and a unique maximum at $t=e^{\beta}(1-\lambda\sigma)^{\sigma/\lambda}$. Above the *ammag* distribution ($\lambda>1/\sigma$) the limit of the density as $t\to 0^+$ is infinity. Only the *ammag* density has a finite, positive limit as $t\to 0^+$.

The curve in Figure 1 representing the inverse *ammag* distribution is also the boundary for the existence of the mean [4]. Indeed, $E(T)<\infty$ for $T\sim \mathrm{GG}(\beta,\sigma,\lambda)$ if and only if $\lambda> -1/\sigma$. Similarly $\lambda> -0.5/\sigma$ is required for a finite variance.

2.3. Relative times and regression models

In many instances, including our application, the primary interest is the comparison of the survival experience of two or more groups of individuals. A classical measure of association is the relative hazard, but an alternative and perhaps more effective way to communicate inferences is to use relative times. Given an unexposed or reference population with survival function $S_0(t)$ and an exposed or treated population with survival function $S_1(t)$, the relative times are defined for $0<p<1$ as the ratio of the corresponding quantile functions, $\mathrm{RT}(p)=t_1(p)/t_0(p)=S_1^{-1}(1-p)/S_0^{-1}(1-p)$

DOI: 10.1002/sim

where S_1^{-1} and S_0^{-1} are the inverses of the survival functions. Strictly speaking, RT(p) should be referred to as 'relative quantiles of survival time' but we have adopted the simpler term, 'relative times'.

The interpretation of RT(p) is that the time required for p per cent of individuals in the exposed or treated population to experience the event of interest is RT(p)-fold the time for the same proportion of events to occur in the reference population. If $(\beta_0, \sigma_0, \lambda_0)$ and $(\beta_1, \sigma_1, \lambda_1)$ denote two different sets of GG parameter values, then $\log \mathrm{RT}(p) = (\beta_1 - \beta_0) + \sigma_1 g_{\lambda_1}(p) - \sigma_0 g_{\lambda_0}(p)$, which reduces to $(\beta_1 - \beta_0) + (\sigma_1 - \sigma_0) g_\lambda(p)$ if $\lambda_1 = \lambda_0 \equiv \lambda$; and RT($p$) is independent of p (i.e. proportional times) only if $\sigma_1 = \sigma_0$.

The conventional (AFT) GG regression model corresponds to GG($\beta' x, \sigma, \lambda$), describing the distribution of the times for the group of individuals with covariate vector x. In this case, the relative times of the group with covariates x to the group with covariates x^* are constant (i.e. proportional times) and are equal to $\exp(\beta'(x - x^*))$. Lack of proportionality of times can be easily incorporated by allowing σ to depend on covariates z, with relative times given by $\exp(\beta'(x - x^*) + \sigma'(z - z^*) g_\lambda(p))$. Full generalization of the model can be accomplished by further allowing λ to depend on covariates. Furthermore, relationships between parameters (e.g. $\lambda = \sigma$) specifying different subfamilies (e.g. gamma) are also possible. Indeed, there may be instances when for certain exposures the times are gamma distributed, while for other exposures the times are *ammag*. Proportionality of hazard functions in the GG family is limited to the particular case of conventional Weibull regression models, which involve only the location parameter.

For the conventional regression model GG($\alpha + \beta' x, \sigma, \lambda$), the form of equation (4) lends itself to a direct measure of the variance of the survival times on the log scale explained by the constellation of variables, x. Specifically, a GG-based R^2 is given by $\beta' \mathrm{Cov}(X) \beta [\beta' \mathrm{Cov}(X) \beta + \sigma^2 V_\lambda]^{-1}$, where V_λ is the variance of the log of the standard GG(0, 1, λ). V_λ is equal to λ^{-2} times the variance of the log of the gamma $G(0, |\lambda|)$ whose moments have been previously described [5].

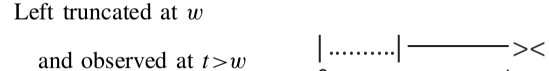
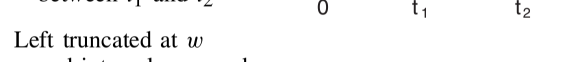
2.4. Estimation of parameters

Since the GG model is fully parametric, the standard approach for estimation is to use the likelihood. For survival data, the form of the likelihood depends on the information provided by the observation. Table I summarizes the most common data types, involving different combinations of left truncation and both right and interval censoring. Software for analysis of data subject to incompleteness due to left truncation and several types of censoring is readily available (e.g., streg of Stata, lifereg of SAS and censorReg of S-Plus). These programs accommodate only AFT regression models for the location parameter, except for Stata, which has options allowing modeling of the scale and shape parameters. SAS can fit conventional regression models for the GG, while S-Plus will not actually fit the full, three parameter distribution. Interestingly, although both the log normal and Weibull distributions are available in all three, the standard two-parameter gamma is not available in any.

Since we wanted to consider regression models involving all three parameters, and wished to be able to fit and compare all of the two-parameter subfamilies of the GG as well, we used SAS PROC NLMIXED, which can handle a general parametric likelihood. The various contributions to the likelihood were easy to program using supplied functions for the gamma CDF and PDF. The NLMIXED procedure also has the ability to compute general functions of the estimated parameters, and a separate feature that allows functions that also depend on time or the covariates. In both cases standard errors are computed by the delta method. Using a supplied function for the gamma quantile function we computed log percentiles for a range of values of p, and values of

DOI: 10.1002/sim

Table I. Patterns of censoring and truncation occurring in survival data, the corresponding likelihood contributions, and capabilities of each of the three most commonly used statistical packages for parametric survival analysis.

Data type		Likelihood contribution	Software features		
			Stata Streg	SAS LIFEREG	S-Plus CensorReg
Observed at t		$f(t) = P(T \in \mathrm{d}t)$	Yes	Yes	Partial*
Right censored at t		$S(t) = P(T>t)$	Yes	Yes	Partial*
Left truncated at w and observed at $t>w$		$f(t)/S(w) = P(T \in \mathrm{d}t \mid T>w)$	Yes	No	Partial*
Left truncated at w and right censored at t		$S(t)/S(w) = P(T>t \mid T>w)$	Yes	No	Partial*
Interval censored between t_1 and t_2		$S(t_1) - S(t_2) = P(t_1<T<t_2)$	No	Yes	Partial*
Left truncated at w and interval censored between t_1 and t_2		$[S(t_1) - S(t_2)]/S(w) = P(t_1<T<t_2 \mid T>w)$	No	No	Partial*

*Not for full generalized gamma. Only for two-parameter cases such as log normal, Weibull and inverse Weibull.

DOI: 10.1002/sim

the log hazard for a range of values of t. Logs of relative times and relative hazards were then computed, with another application of the delta method to calculate standard errors. Large sample, pointwise confidence limits on the log scale were transformed to obtain asymmetric confidence bands for relative times and relative hazards. Approximate simultaneous confidence bands may also be obtained [9]. Additional details on using this procedure for model fitting, as well as the computation of relative times and relative hazards, are provided in Section 5.

An alternative approach to assessing the variability of estimated parameters, as well as more complex quantities such as relative times, was implemented using 500 bootstrap samples. Our results include a comparison of 90 per cent confidence intervals based on percentiles of the bootstrap distribution for estimates of parameters, relative times and relative hazards with symmetric (for the relative times and relative hazards, symmetric on the log scale) intervals based on the standard errors produced by the delta method.

3. APPLICATION

3.1. Study design and aims

Individuals infected with HIV are at risk for AIDS and subsequent death. Treatments for HIV have evolved from a period of no available therapy (prior to 1987) to combinations of three or more drugs (after 1995), collectively referred to as highly active antiretroviral therapy (HAART) [10]. Since the evolution of HIV therapy has followed a definite pattern over time, it is possible to define sequential calendar periods corresponding to distinct therapeutic eras. Following Schneider *et al*. [7], we defined four therapy eras, beginning with an initial period of no or only monotherapy (July 1984–December 1989), followed by a period of mono or combination therapy (January 1990–December 1994), then by the introduction of HAART (January 1995–June 1998) and finally by the (short to moderate-term stable) era of HAART (July 1998–December 2003). This approach allows assessment of the changing pattern of survival at the population level after the development of clinical AIDS from 1984 through the HAART eras [7].

We based our analysis on data from two large cohort studies, which have longitudinally followed HIV-infected individuals. The first is the Multicenter AIDS Cohort Study (MACS), an ongoing study of HIV-1 infection in a cohort of homosexual and bisexual men, which began in 1983 [11]. The second is the Women's Interagency HIV Study (WIHS), which is also a multicenter, prospective cohort study begun in 1994 [12]. Both studies have used similar methods and the same co-ordinating centre. Both studies conduct semi-annual interviews, which include detailed questions about the use of antiretroviral therapy. In both studies deaths are ascertained using both active and passive methods, including abstraction of death certificates and searches of national death registries. Data from both the MACS and WIHS show that a high percentage of participants with clinical AIDS in both studies were actually using the indicated therapy during the corresponding therapeutic era [7].

The analysis included a total of 1504 men and 461 women with follow-up after an incident diagnosis of clinical AIDS. In this study, AIDS was defined using the 1993 CDC surveillance criteria [13] based on clinical conditions (i.e. excluding the laboratory criterion of low CD4 count). Individuals with an incident clinical AIDS diagnosis in either of the first two periods who were still alive at the end of the period were treated in the analysis as censored observations at the end of the period in which the incident diagnosis occurred; they did not contribute data to

DOI: 10.1002/sim

subsequent calendar periods. This was done to ensure comparability between the men and women in the final two periods [7] since data from the WIHS was not available until January 1995. Individuals with an incident AIDS diagnosis in the third period who were alive at the end of the period were censored at the end of the third period and treated as left truncated observations in the fourth period. Furthermore, for WIHS participants (periods 3 and 4), the date of the AIDS diagnosis could only be determined as the midpoint between two consecutive visits, so half the length of this interval was included in the analysis as left truncation. Calendar periods are external, time-dependent covariates since they change equally over time for different individuals.

3.2. Results

Of the combined sample of 1965 HIV infected individuals observed to develop AIDS from July 1984 to December 2003, the numbers with follow-up in the no/mono therapy, mono/combination therapy, HAART introduction and stable use of HAART eras were 633, 660, 472 and 549, respectively. The corresponding numbers of deaths in the four eras were 388 (61 per cent), 445 (67 per cent), 109 (23 per cent) and 111 (20 per cent), respectively. The fourth era included 200 individuals who developed AIDS after the beginning of the period (July 1998) and 349 individuals who developed AIDS during the third era and thus contributed left-truncated observations to the fourth era.

For each era, Figure 3 depicts the logarithm of the cumulative hazard functions from the non-parametric product-limit estimate and the corresponding fitted GG distribution, using the saturated model where each period has its own (β, σ, λ) parameters for a total of 12 parameters. In each case,

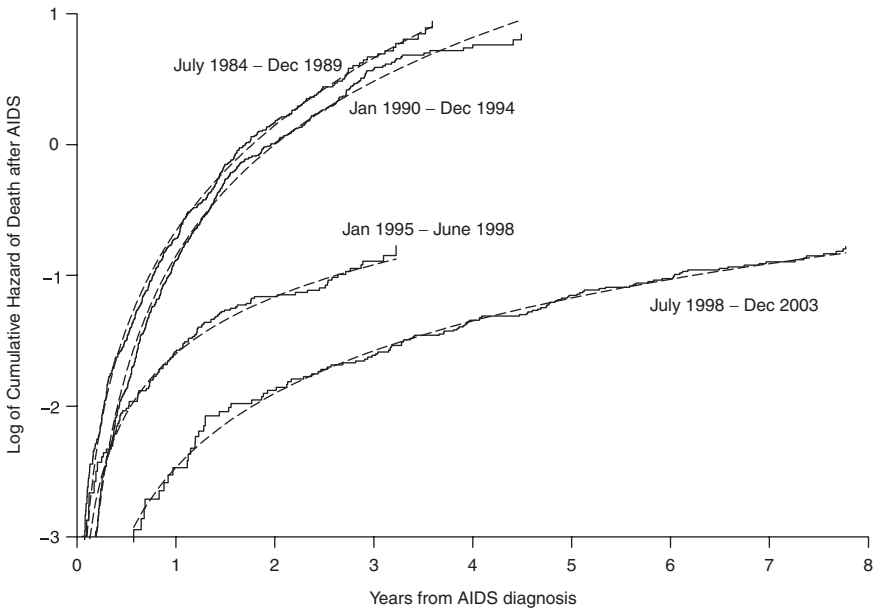


Figure 3. Non-parametric and GG-based estimates of the log of the cumulative hazard of death after an incident AIDS diagnosis in each of the four therapeutic eras for HIV.

DOI: 10.1002/sim

the fit was clearly very good; the only signs of lack of fit occurred approximately 3 years after AIDS in period two, where the non-parametric curve indicated relatively few events. The violation of proportional hazards was evident as the distances between the curves were not constant over time, with the possible exception of periods 3 and 4.

Table II shows a series of regression models used to determine whether simpler models were as appropriate as the saturated one, and to compare inferences with the semi-parametric proportional hazards model. The saturated model has 12 parameters (i.e. three parameters (β, σ, λ) for each of the four periods); conventional AFT regression models allow only the beta parameter to depend on covariates while the other two parameters are common or fixed. The conventional Weibull regression model corresponding to $\mathrm{GG}(\beta' x, \sigma, 1)$ provided practically the same relative hazards as the Cox proportional hazards model, suggesting no advantage to leaving the reference hazard as arbitrary in this analysis. Since the Weibull model satisfies the assumption of proportional times, Table II includes the relative times as well, whereby the survival times in periods two, three and four were estimated to be 1.17, 3.80 and 10.07 fold longer than the corresponding times in period one. Next, we considered the conventional GG regression model corresponding to $\mathrm{GG}(\beta' x, \sigma, \lambda)$. In this case the relative hazards are time-dependent, but the relative times are constant. Although the relative times under the conventional GG model were similar to the relative times for the Weibull model, the GG model fit significantly better, both by the likelihood ratio test and by the Wald test for λ.

The saturated GG model corresponding to $\mathrm{GG}(\beta' x, \sigma' x, \lambda' x)$ fit significantly better than the conventional GG model, providing evidence that both the relative hazards and the relative times were not constant (i.e. the scale and/or shape parameters were heterogeneous). For the final model, we determined whether any of the two-parameter subfamilies (Weibull, log normal, gamma and *ammag* distributions) provided a more parsimonious fit to the data. The final column of Table II shows the parameter estimates and standard errors for the best fitting model in each of the four therapy eras corresponding to the eight parameter model $\mathrm{GG}(\beta' x, \sigma' x, (1/\sigma_1, \sigma_2, 1/\sigma_3, 1))$ (i.e. *ammag*, gamma, *ammag* and Weibull). The likelihood ratio statistic with four degrees of freedom comparing the saturated and final models was very small (0.6), providing support for the final model. In each of the four periods, the estimates of the scale and shape parameters for the saturated GG model were consistent with the restriction defining the best fitting two-parameter model (e.g. for period one, $\hat{\lambda} = 1.369 \approx 1.293 = \frac{1}{0.773} = 1/\hat{\sigma}$). The standard errors for the scale parameters in the last two periods (σ_3 and σ_4) are considerably smaller for the two-parameter models than for the full GG, suggesting that the full model was over-fitting the data in these two periods.

The left-hand panel of Figure 4 displays the parameter estimates for the model providing the best fit in each of the four periods (see final column in Table II) in the (σ, λ) half-plane, together with confidence intervals based on 5th and 95th percentiles of the bootstrap distribution. The estimates of the time corresponding to 25 per cent mortality ($Q1$) were included to provide a measure of location (i.e. β). Using open symbols, we also show the corresponding parameter values for the saturated GG model. Except for period four, the locations of the open and closed (for the final model) symbols were very close. The figure clearly illustrates changes in the distribution of survival times after a diagnosis of clinical AIDS in the four therapy eras. Furthermore, in each of the first three periods the confidence intervals exclude a hazard with a different shape from that of the best fitting distribution. The confidence interval for period four includes the exponential distribution, and we cannot entirely rule out the possibility of a distribution with increasing hazard ($\sigma < 1 = \lambda$), consistent with the possible emergence of resistance to HAART during this final period. The symmetric interval for σ was (0.89, 1.70) and the asymmetric interval was (0.95, 1.76), compared to (0.95, 1.80) for the bootstrap.

DOI: 10.1002/sim

Table II. Parameter estimates and standard errors (SE) for regression models for survival after AIDS in four eras of therapy for HIV. Also included are likelihood ratio statistics (LRS) for nested models, and both relative hazards and relative times in the case of proportionality.

Period (Therapy era)	Parameter		Semi-parametric Cox	Parametric models: Weibull conventional	Generalized gamma, conventional	Generalized gamma, saturated	Generalized gamma, final
1: July 1984–December 1989	Location	β_1	0 (ref)	0.553(0.046)	0.643(0.055)	0.674(0.072)	0.637(0.051)
2: January 1990–December 1994		β_2	−0.165(0.070)	0.707(0.043)	0.794(0.053)	0.649(0.060)	0.648(0.041)
3: January 1995–June 1998		β_3	−1.500(0.110)	1.888(0.090)	1.961(0.097)	2.446(0.247)	2.432(0.211)
4: July 1998–December 2003		β_4	−2.379(0.148)	2.863(0.087)	2.914(0.091)	3.071(0.483)	3.106(0.174)
1	Scale	σ_1	NA	0.901(0.024)	0.804(0.042)	0.773(0.068)	0.798(0.061)
2		σ_2	NA	σ_1	σ_1	0.847(0.048)	0.848(0.024)
3		σ_3	NA	σ_1	σ_1	2.008(0.803)	2.082(0.300)
4		σ_4	NA	σ_1	σ_1	1.376(1.036)	1.293(0.245)
1	Shape	λ_1	NA	1	1.270(0.110)	1.369(0.189)	$1/\sigma_1$
2		λ_2	NA	1	λ_1	0.851(0.127)	σ_2
3		λ_3	NA	1	λ_1	0.549(0.697)	$1/\sigma_3$
4		λ_4	NA	1	λ_1	0.845(1.879)	1
	−2 log likelihood		14029.5	4545.6	4538.7	4501.5	4502.1
	LRS; d.f.			6.9; 1	37.2; 6	0.6; 4	
Relative hazard							
Period 2 to Period 1			0.85 (0.06)	0.84 (0.06)	Non-constant*	Non-constant	Non-constant
Period 3 to Period 1			0.22 (0.02)	0.23 (0.02)	Non-constant	Non-constant	Non-constant
Period 4 to Period 1			0.09 (0.01)	0.08 (0.01)	Non-constant	Non-constant	Non-constant
Relative time							
Period 2 to Period 1			†	1.17 (0.07)	1.16 (0.07)	Non-constant*	Non-constant
Period 3 to Period 1			†	3.80 (0.38)	3.74 (0.38)	Non-constant	Non-constant
Period 4 to Period 1			†	10.07 (0.99)	9.69 (0.96)	Non-constant	Non-constant

*Non-constant = time-dependent in the case of relative hazards, or percentage-dependent in the case of relative times.

†Measure not consonant with baseline hazard being arbitrary.

DOI: 10.1002/sim

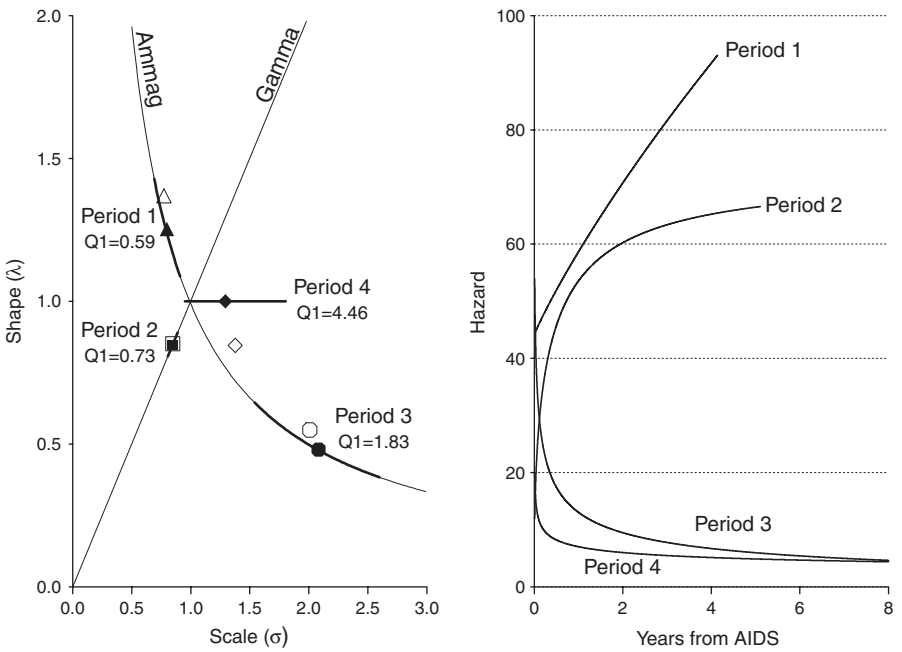


Figure 4. Left panel: best fitting two-parameter GG model for each of the four periods, with percentile-based bootstrap confidence intervals. Open symbols represent the parameter estimates for the saturated GG model. Right panel: estimated hazard function for the best fitting two-parameter distribution in each of the four periods.

The right-hand panel of Figure 4 depicts the hazard functions of the two-parameter models providing the best fit in each of the four periods (see final column in Table II). In period one the *ammag* hazard increased steadily from a finite level after the initial diagnosis of AIDS. In contrast, the gamma hazard for the second period was lower initially, and although it approached a horizontal asymptote instead of increasing indefinitely, the cumulative mortality was not substantially different from that in period one. Specifically, the hazard functions for the first two periods are plotted up to the 95th percentiles, which were 4.2 years for period one and 5.1 years for period two. In period three we have an *ammag* hazard with a decreasing failure rate from a finite hazard at time zero. There was a large reduction in the hazard between periods two and three. The Weibull hazard in period four was also monotone decreasing, and it did not exceed that for period three before 8 years following an AIDS diagnosis, indicating that the effectiveness of HAART was maintained if not improved as the use of HAART became high and stable.

To more directly compare the changes in survival during the four therapeutic eras we consider the relative times (with period one as the reference) using the best fitting model for each of the four periods. These are shown in Figure 5, with pointwise 90 per cent confidence bands derived from the delta method. Also shown are boxplots based on 500 bootstrap samples for $p = 10$, 20 and 30 per cent. The relative times for period two *vs* period one indicated only modest improvement in survival. The greatest benefit of therapy during the second period was clearly to improve initial survival. In sharp contrast, the introduction of HAART in period three resulted in less improvement

DOI: 10.1002/sim

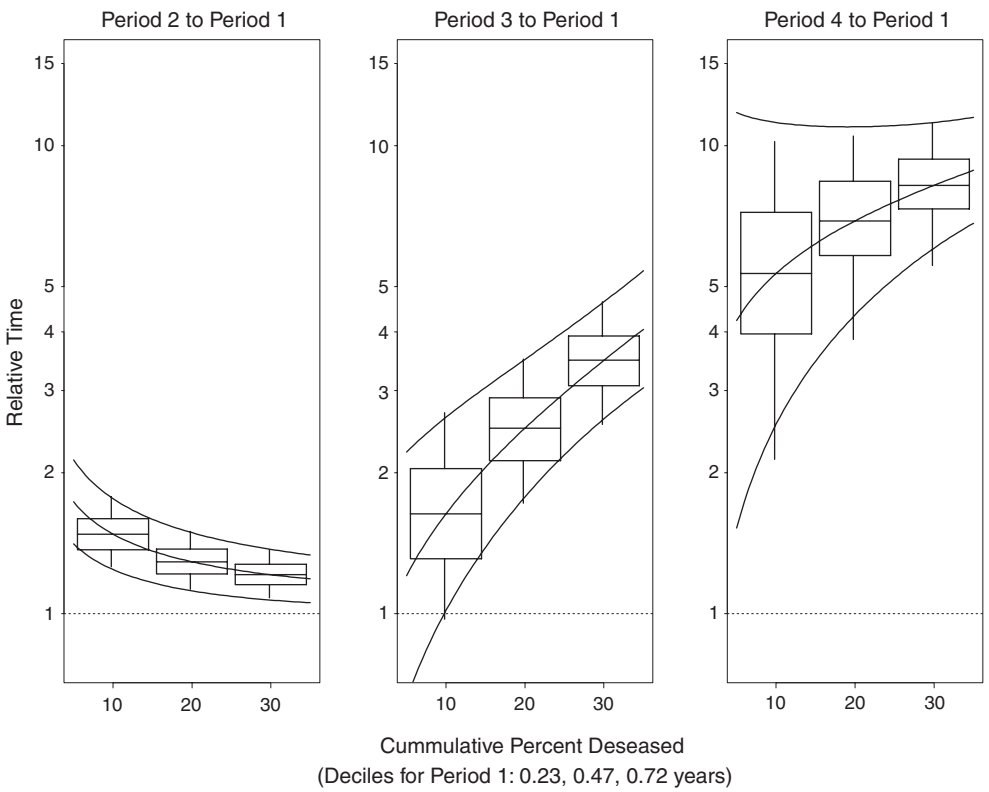


Figure 5. Relative times of periods 2, 3 and 4 to period 1 for the best fitting two-parameter GG distributions. Ninety per cent confidence bands based on the delta method. Boxplots for $p = 10, 20$ and 30 per cent based on 500 bootstrap samples, with midline representing the fitted values and whiskers indicating the 5th and 95th percentiles.

initially, but considerably improved survival with the passage of time. Finally with the stable use of HAART, the improvement in survival was evident throughout the entire period of observation.

The same therapeutic effects are shown in Figure 6 in terms of relative risk in the plot of the relative hazards (again using period one as the reference, with similar confidence bands and boxplots). This figure illustrates the lack of proportionality of the hazard functions for the four periods. Only in the final period was there any suggestion of flatness in the relative hazard, which would indicate approximate proportionality. The comparison of the first panel with the final two panels again provides a clear picture of the benefits of HAART for the treatment of clinical AIDS.

4. DISCUSSION

Our results demonstrate the changing therapeutic panorama for AIDS and provide an illustration of hazard functions that are not proportional. For these kinds of data, which include both censoring and truncation, parametric survival models offer a practical, useful approach to inference. The GG

DOI: 10.1002/sim

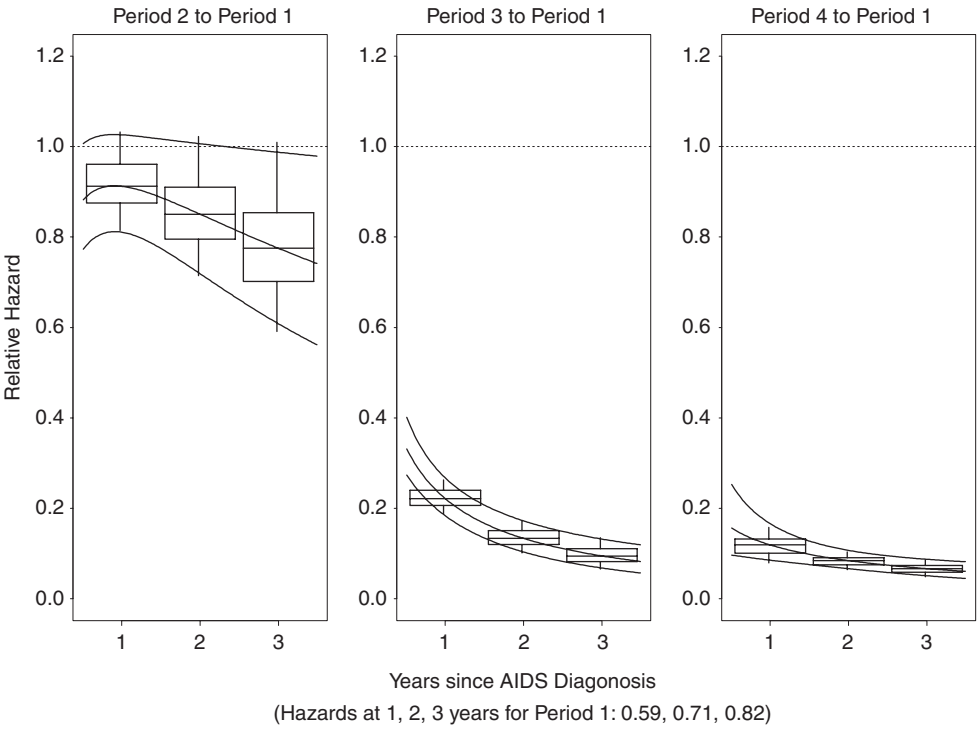


Figure 6. Relative hazards of periods 2, 3 and 4 to period 1 for the best fitting two-parameter GG distributions. Ninety per cent confidence bands based on the delta method. Boxplots for $t = 1, 2$ and 3 years based on 500 bootstrap samples, with midline representing the fitted values and whiskers indicating the 5th and 95th percentiles.

distribution is a broad family that includes most of the commonly used parametric distributions, allowing comparisons among the different special cases. Although it does not contain the commonly used log logistic family, this distribution is well approximated by a log normal with parameters β and 1.82σ. Nonetheless, the family may not have members to describe the survival function of some diseases or patterns of mortality.

Using the time to death after an incident diagnosis of clinical AIDS in four therapeutic eras, we have shown that the Cox model limits the inference when the proportional hazards assumption is violated. The GG approach provided additional insight into the behaviour of the hazard functions in the four therapeutic eras. Future simulation studies may shed some additional light on the failure of the Cox proportional hazards model when the true underlying distribution is indeed the GG, and, conversely, the potential inefficiency of the GG when the underlying hazards are indeed proportional.

The GG family includes the four common types of hazard function (increasing and decreasing failure rates, bathtub and arc-shaped hazards), as graphically depicted by the taxonomy of Figure 1 and illustrated in Figure 2. This makes the GG distribution particularly useful since the shape of the hazard function may provide additional clues in model selection. For example, an increasing hazard might be expected in a study of a disease caused by an infection with no host response,

DOI: 10.1002/sim

whereas an arc-shaped hazard might result when either a host response occurs or an effective treatment is available. An intervention having sustained effectiveness might produce a decreasing hazard, while one with transient effectiveness due to emergence of resistance might have a hazard function with a bathtub shape.

The GG is one of a limited number of parametric distributions to include a bathtub-shaped hazard. Other distributions include the exponential power distribution of Smith and Bain [14], whose only other possible hazard is increasing. A second family is the three-parameter, generalized Weibull distribution of Mudholkar *et al.* [15] which includes all four common types of hazard, and in addition is closed under the proportional hazards relationship unlike the GG distribution. However, generalized Weibull distributions with bathtub and increasing failure rates have finite support, with an upper limit depending on one or more of the parameters, depending on the parameterization. For this reason, standard large sample theory does not apply and alternative methods of parameter estimation must be employed for these members of the generalized Weibull family.

A limitation of the graphical depiction provided in Figure 1 is that it is based on the scale and shape parameters only. These parameters determine the shape of the hazard function, but not its overall level. To complement the taxonomy for the application in Figure 4, we also included the estimates of the first quartiles to provide a measure of location and to convey a sense of the overall levels of the hazard functions. This is useful because two points in Figure 1 close to each other may correspond to different levels of the hazards functions depending on the values of the location parameters. Another limitation is that difficulties in distinguishing different members of the GG family have been noted previously [5], and indeed in period four there is some degree of uncertainty about the shape of the hazard function at very early event times. One reason for this is the truncation introduced by uncertainty in the exact timing of the initial AIDS diagnosis, caused in part by missed visits that were more prevalent in the WIHS than in the MACS.

The GG distribution is actually included in a larger family, the four-parameter generalized F distribution [16], for which the conventional AFT model can be fitted by the S-Plus GFCURE package [17]. In this case the random variable $(e^{-\beta}T)^{1/\sigma} \sim F(2m_1, 2m_2)$, an F distribution with non-integer degrees of freedom ($m_1>0, m_2>0$). With this parameterization the GG family corresponds to two different limiting cases, $m_1 \to \infty$ ($\lambda<0$) and $m_2 \to \infty$ ($\lambda>0$). The generalized F distribution was applied in the survival context by Ciampi *et al*. [18]. These authors noted, however, that this is an extremely rich family, having members that resemble each other very closely, with the result that the optimization required for estimation of the parameters is often ill-conditioned. For this reason, and because of the flexible properties of the GG family that we have already discussed, we did not consider the more extended family.

Several approaches to the analysis of survival data with non-proportional hazards have been proposed which are not fully parametric. These range from the standard inclusion in Cox regression models of exposure by time interactions, to the use of cubic splines to smooth the baseline hazard [19, 20], to the more complete use of time-varying covariates [21]. An additional motivation for the development of these approaches has been the limited number of types of hazard available using the standard distributions, in particular the lack of a bathtub-shaped hazard [20]. We have discussed an extensive, fully parametric family of distributions with sufficient flexibility to incorporate the four common types of hazards, and with the particular advantage of providing relative times to quantify the expansion or contraction that exposures exert on survival.

An important distribution for the taxonomy of the hazard, the *ammag*, has not been previously considered, but appears to provide an additional parametric family that may be useful for other studies with survival as an outcome. In our example, which included a variety of different types

DOI: 10.1002/sim

of hazard functions, the *ammag* distribution provided the best fit for two of the four periods. This distribution is the only member of the GG family having a finite, non-zero hazard at time zero. Among commonly used parametric distributions, only the Gompertz has this property [22].

Overall, we conclude that the GG distribution provides a flexible and practically useful alternative to the Cox proportional hazards model. The taxonomy provided in Figure 1 illustrates the behaviour of the hazard corresponding to different special cases, and it is hoped will encourage wider use of the GG family.

5. STATISTICAL SOFTWARE

We describe here statistical software for the use of the generalized gamma distribution as a framework for parametric survival analysis. We first present the essential commands for Stata 9, SAS 9.1 and S-Plus 7 for conventional regression models and indicate their limitations. To overcome the limitations of currently available software (e.g. to fit an *ammag* model), we discuss the use of SAS PROC NLMIXED for model fitting and subsequent computation of relative times and relative hazards, with standard errors computed by the delta method. Corresponding functions in S-Plus and R, as well as the full data for the application along with the Stata, SAS and S-Plus programs to reproduce the analyses can be downloaded from the web site: http://statepi.jhsph.edu/software.

5.1. Software in Stata, SAS and S-plus for parametric models

The data types that can be handled by Stata, SAS and S-plus were described in Table I. The major limitation of the Stata streg command is the inability to handle interval censored observations. If the widths of the censoring intervals are relatively narrow, then interval censored observations may be treated as events occurring at the mid-point of the interval. The major limitation of the SAS lifereg procedure is the inability to handle left truncated observations. Only in cases where the entry times are relatively small (i.e. well before events start to occur) can left truncation be ignored. Although the S-plus censorReg function can handle all types of censoring and truncation, it will only fit special cases of the GG. However, the user contributed S-Plus packages GFCURE [17] and ACCFLF (freely downloadable from StatLib at http://lib.stat.cmu.edu/) can be used to fit the GG AFT model.

The key commands to implement conventional GG (log normal for S-Plus) regression are shown in Table III. In these commands, the time until event or censoring is the variable *exit* (>0), the entry time for left truncated observations is *entry* ($\geqslant 0$), and the censoring variable is *event* ($0 =$ censored, $1 =$ event). Replacing *exit*event(0)* by *(left, right)* provides the means to handle interval censored data in SAS. Substituting `weibull` or `lnormal` for `gamma` in Stata and SAS; and `weibull` for `lognormal` in S-Plus will fit Weibull and log normal models, respectively. For non-conventional GG regression models, Stata has the options anc and anc2 which allow separate models for the logarithm of the scale parameter and the shape parameter, respectively (Table III). For additional discussion of the application of parametric models in survival analysis using Stata and SAS see [23, 24], respectively.

The following listing contains selected observations from the data set used in the application. The listing includes each of the data types occurring in the four eras of therapy. Left truncation occurs only in the third and fourth periods. The variables `p1`–`p4` are indicator variables for the four periods and the variable `period` takes values 1 to 4. In these data the entry and exit times

DOI: 10.1002/sim

Table III. Basic Stata, SAS, and S-PLUS programs for parametric conventional regression analyses for time-to-event data with right censoring and left truncation.

Software	Commands
Stata*	`stset` *exit*, `failure(`*event*`) enter(`*entry*`)` `streg` *covariate list*, `distribution(gamma)`
SAS**	`proc lifereg;` `model` *exit*`*`*event*`(0) =`*covariate list*`/distribution=gamma;`
S-Plus***	`censorReg(censor(`*exit*`,`*event*`)~`*covariate list*`,` `distribution='lognormal', truncation=censor(`*entry*`, NA, 2))`

*The options anc(*covariate list*) and anc2(*covariate list*) allow separate regression models for the log scale and shape parameters, respectively.
**Left truncation coded in the variable *entry* is not permitted in SAS.
***Generalized gamma not available.

are measured in years. Note that the final two records in the listing were contributed by the same individual, who entered period three after an initial AIDS diagnosis (late entry), was censored at the end of period three at 1.335 years after diagnosis, then entered period four (late) and was finally censored at the end of period four.

publicID	entry	exit	event	p1	p2	p3	p4	period
5083	0.000	1.914	1	1	0	0	0	1
3500	0.000	4.410	1	0	1	0	0	2
2113	0.000	3.239	0	0	0	1	0	3
75560	0.500	5.369	0	0	0	0	1	4
99730	0.261	1.335	0	0	0	1	0	3
99730	1.335	6.835	0	0	0	0	1	4

5.2. *Software in Stata for the application in Section 3*

The following edited listing is from the Stata commands (in bold) fitting the full, three-parameter distribution to the data from period one.

```
. stset exit, failure(event) enter(entry)
. streg if p1 == 1, distribution(gamma)

Log likelihood = -783.8584
```

_t	Coef.	Std. Err.	(...)
_cons	.6736712	.0716252	(...)
/ln_sig	-.2572981	.0878631	(...)
/kappa	1.368894	.1890939	(...)
sigma	.7731377	.0679303	(...)

The values of `Coef.` and `Std. Err.` for `_cons`, `sigma` and `/kappa` correspond to β_1, σ_1 and λ_1 in the saturated GG model for period 1 in Table II. The coefficient `/ln_sig` is the logarithm of sigma with a standard error obtained by the delta method. The values of the log likelihood for

DOI: 10.1002/sim

periods 2, 3 and 4 were −802.16, −358.51 and −306.24. The sum for all four periods equals −2250.77, giving a value for −2 log likelihood of 4501.5 for the saturated model as shown in Table II.

The following results are from a Stata command to fit the six-parameter conventional GG model to the full data from all four periods.

```
. streg p1 p2 p3 p4, distribution(gamma) noconstant
* Include the noconstant option because there is no
intercept, i.e. no reference category.

Log likelihood = -2269.3491
```

_t	Coef.	Std. Err.	(...)
p1	.642839	.0552769	
p2	.7937214	.0528871	(...)
p3	1.961447	.096863	(...)
p4	2.914392	.090706	(...)
/ln_sig	-.217642	.0524492	(...)
/kappa	1.270146	.109724	(...)
sigma	.8044134	.0421908	(...)

The `Coef.` and `Std. Err.` values for `p1`–`p4`, `sigma` and /`kappa` correspond to those for β_1 to β_4, σ_1 and λ_1 in the GG AFT model in Table II. Furthermore, the −2 log likelihood gives the value 4538.7 shown in Table II. In GG AFT regression models the relative times are constant and may be obtained using Stata by eliminating one of the dummy variables (e.g. `p1`, making period one the reference category) and using the `tr` option. The saturated GG model can be fitted using the single command **`streg p2 p3 p4, d(gamma) anc(p2 p3 p4) anc2(p2 p3 p4)`**.

The conventional AFT Weibull model can also be easily obtained using Stata; a similar option, `hr`, provides the relative hazards. With the exception of the Weibull model for period four however, Stata cannot directly provide the final GG models (i.e. gamma and *ammag*). This limitation motivates the use of more general software which we provide in the next subsection.

5.3. Software in SAS (PROC NLMIXED) for the application in Section 3

In order to compute the log likelihood, relative hazards and relative times when using the general purpose program PROC NLMIXED of SAS, it is necessary to compute both the CDF and PDF, as well as the quantile function of the GG distribution GG(β, σ, λ). For this purpose we used supplied functions for the standard two-parameter gamma distribution, density and quantile functions. The parameterization used by these functions is slightly different from that we have used for $G(\beta, \sigma)$. For our parameterization of the two-parameter gamma, the syntax is *function*('*gamma*', arg, σ^{-2}, $\sigma^2 e^{\beta}$), where the generic term *function*, may be either *cdf*, *pdf* or *quantile*, and the generic (second) argument denoted by arg is either time ($t>0$) for *cdf* and *pdf*, or proportion ($0<p<1$) for *quantile*.

As indicated at the end of Section 2.1.1, the GG survival function is then

$$S_{\mathrm{GG}(\beta,\sigma,\lambda)}(t) = S_{\mathrm{G}(\beta\lambda/\sigma,|\lambda|)}(t^{\lambda/\sigma}) = 1 - cdf(\text{'}gamma\text{'}, t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma}) \quad \text{if } \lambda>0$$

$$= cdf(\text{'}gamma\text{'}, t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma}) \quad \text{if } \lambda<0$$

DOI: 10.1002/sim

The density function is obtained by differentiating the survival function with respect to time

$$f_{\mathrm{GG}(\beta,\sigma,\lambda)}(t) = \left| \frac{\mathrm{d}t^{\lambda/\sigma}}{\mathrm{d}t} \right| f_{G(\beta\lambda/\sigma,|\lambda|)}(t^{\lambda/\sigma}) = \frac{|\lambda|}{\sigma t} t^{\lambda/\sigma} pdf(\text{`}gamma\text{'}, t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 \mathrm{e}^{\beta\lambda/\sigma})$$

Alternatively, the coding for these functions could be based on the AFT model as indicated in equation (1).

These expressions can be used to compute the contribution to the likelihood for each observation. For example, for the case of an event with late entry (data type 3 in Table I), the following code calculates the likelihood represented by `ll`.

```
/* Log likelihood for uncensored observations, with late entry:
  log[f(t)t/S(w)] */
if (entry > 0 and event = 1 and lambda < 0) then
    ll = log(pdf('gamma', exit**(lambda/sigma),1/(lambda*lambda),
      lambda*lambda*exp(beta*lambda/sigma)))
        + lambda*log(exit)/sigma + log(-lambda) - log(sigma)
  - log( cdf('gamma',entry**(lambda/sigma),1/*lambda*lambda),
    lambda*lambda*exp(beta*lambda/sigma)));
if (entry > 0 and event = 1 and lambda > 0) then
  ll = log(pdf('gamma', exit**(lambda/sigma),1/(lambda*lambda),
    lambda*lambda*exp(beta*lambda/sigma)))
      + lambda*log(exit)/sigma + log(lambda) - log(sigma)
  -log(1-cdf('gamma',entry**(lambda/sigma),1/(lambda*lambda),
    lambda*lambda*exp(beta*lambda/sigma)));
```

Since $-\log(t)$ does not depend on the parameters, we exclude this term from the log of the PDF as done by both SAS and Stata. To avoid large arguments for the supplied functions that can result in numerical problems, the time variable should be appropriately scaled (e.g. from days to years). The key commands in NLMIXED to obtain the MLE are:

```
proc nlmixed fd;
  * Initial values for the parameters;
  parms beta=0 sigma=1 lambda=0.5;
  * Compute ll, the log likelihood for the current observation;
  model exit ~ general(ll);
run;
```

The keyword `general` tells the program that the variable `ll` is the value of log likelihood function for the current observation. For this keyword, the variable to the left of the tilde is simply a place holder, and plays no role in the optimization. For this application we used numerical derivatives (the `fd` option in the `nlmixed` statement), which gave accurate results even for the complex functions involved. The scale parameter must of course be positive. This constraint can be enforced using an additional program statement (`bounds sigma>0;`), or through a log linear model for this parameter (parameter `logsig` and an extra program statement, `sigma = exp(logsig);`); we did not need either of these options in our application.

DOI: 10.1002/sim

The initial values in this sample program are for the standard GG(0, 1, 0.5); in our application we did not find that any refinement was needed, although this may not always be the case. One way to provide better initial values would be to use summary statistics such as the median and IQr; another would be to use one of the standard packages to fit an initial exponential or Weibull model. Alternatively one can perform an initial search by including ranges of values for the parameters in the `parms` statement.

The NLMIXED procedure also has the ability to compute general functions of the estimated parameters, with standard errors computed by the delta method. For example, the statement `estimate 'log sigma' log(sigma);` estimates the log of the scale parameter, which may be more normally distributed. In addition, the procedure has a separate feature that allows functions that also depend on time or covariates, with standard errors again computed by the delta method. For example, as shown in Section 2.1.2, the GG quantile function is as follows:

$$t_{\mathrm{GG}(\beta,\sigma,\lambda)}(p) = [quantile(\text{'}gamma\text{'}, p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \quad \text{if } \lambda > 0$$

$$= [quantile(\text{'}gamma\text{'}, 1-p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \quad \text{if } \lambda < 0$$

Using the supplied gamma quantile function, we computed log percentiles for a range of values of p ($0<p<1$), and values of the log hazard for a range of $t>0$ values. To do the former we created an additional file of equally spaced values of p, with the desired range and number of points (e.g. 100) sufficient for plotting purposes and dummy values for the other variables. This file was then appended to the original data file. In addition, a `weight` variable was defined to be one for each of the original observations and zero for the additional data, and used with the `replicate` statement in the procedure to exclude the additional observations from the fitting process. Additional statements were added to compute the log of the pth percentile of the fitted GG distribution for each value of p in the appended data set. These quantiles were included in a `predict` statement, which put them (`Pred`) in an output file (`results`) together with the standard errors (`StdErrPred`). To compute the log of the hazard for each of a set of appended values of *time*, we first computed both the density and the survival function, which are also saved in the output data set using the `id` statement. Note that a similar program can produce confidence bands for the survival function.

As previously indicated, the primary motivation for the use of NLMIXED is to be able to fit models not available in the standard statistical packages. Specifically, we can fit the two-parameter gamma and the *ammag* among many others, as well as allow regression in both the scale and shape parameters. Since for computational reasons the case $\lambda = 0$ is best handled separately, it seems best to use the scale (σ) parameter for restricted models, and then compute a shape (λ) value that is either positive or negative, depending on the desired two-parameter distribution. Thus to fit the two-parameter gamma and *ammag* models, the `lambda` parameter must be removed from the parameter list and the program must contain additional statements of the form `lambda = sigma;` or `lambda = 1/sigma;` respectively. For the special case of the log normal distribution ($\lambda = 0$) the same functions are available for the normal distribution (keyword '`normal`'), and a separate program can be written. Alternatively one can simply fix `lambda` to be a small value (e.g. `lambda = 0.001;`).

The following is a partial listing of a `results` file from the estimation of log percentiles for the *ammag* fit for period one. The file includes the values of the log percentiles (`Pred`) corresponding to the values of p (`prob`), their asymptotic standard errors and 95 per cent Wald-type

DOI: 10.1002/sim

confidence limits. For completeness, the actual percentiles (qp) were also included using the id statement.

Obs (...)	period	weight	prob	qp	Pred	StdErr Pred	Alpha	Lower	Upper
4	1	0	0.025	0.05663	−2.87120	0.077567	0.05	−3.02352	−2.71888
5	1	0	0.030	0.06798	−2.68854	0.077331	0.05	−2.84040	−2.53669
6	1	0	0.035	0.07934	−2.53403	0.077084	0.05	−2.68540	−2.38266
7	1	0	0.040	0.09071	−2.40010	0.076826	0.05	−2.55096	−2.24923

Similarly, the following lines are from the estimation of the log hazard (Pred) for this fit. The values of the density (ft), survival (st) and hazard (per cent) functions have also been included. Note the consistency between line 7 in the previous listing and line 3 in the following.

Obs	period	weight	exit	ft	st	hazard	Pred	StdErr Pred
1	1	0	0.01	0.44182	0.99558	44.3781	−0.81242	0.078565
2	1	0	0.05	0.44096	0.97792	45.0916	−0.79647	0.078220
3	1	0	0.09	0.43955	0.96031	45.7714	−0.78151	0.077308
4	1	0	0.13	0.43772	0.94277	46.4296	−0.76723	0.076102

5.4. Software in S-Plus and extensions of Stata for the application in Section 3

We have developed functions in S-Plus which correspond to those presented for SAS in Section 5.3. As in SAS, S-Plus provides functions for the CDF (pgamma), PDF (dgamma) and quantiles (qgamma) for the two-parameter gamma distribution [25], which as discussed in Section 2 allows the definition of the corresponding functions for the GG. In addition, the general optimizing routine nlminb of S-Plus and optim in R play the role that NLMIXED does in SAS [25]. On the website, http://statepi.jhsph.edu/software we have included an extended version of Section 5, with detailed code in S-Plus to accomplish the analyses presented in Table II. Although we do not consider them, the general optimization routines ml in Stata [26] could also be used to fit all of the models considered here.

ACKNOWLEDGEMENTS

Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS) with centers (Principal Investigators) at The Johns Hopkins University Bloomberg School of Public Health (Joseph B. Margolick, Lisa Jacobson), Howard Brown Health Center and Northwestern University Medical School (John Phair), University of California, Los Angeles (Roger Detels), and University of Pittsburgh (Charles Rinaldo). Additional data were collected by the Women's Interagency HIV Study (WIHS) Collaborative Study Group with centers (Principal Investigators) at New York City/Bronx Consortium (Kathryn Anastos); Brooklyn, NY (Howard Minkoff); Washington DC Metropolitan Consortium (Mary Young); The Connie Wofsy Study Consortium of Northern California (Ruth Greenblatt); Los Angeles County/Southern California Consortium (Alexandra Levine); Chicago Consortium (Mardge Cohen); Data Coordinating Center (Stephen Gange). We thank Stephen Cole for writing *gamma* backwards when we were describing the *ammag* distribution in a research seminar.

Studies were approved by the Committees of Human Research of participating institutions. Websites for both studies and for software in this paper are located at http://statepi.jhsph.edu

DOI: 10.1002/sim

REFERENCES

1. Wei LJ. The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. *Statistics in Medicine* 1992; **11**:1871–1879.
2. Muñoz A, Sunyer J. Comparison of semiparametric and parametric survival models for the analysis of bronchial responsiveness. *American Journal of Respiratory Critical Care Medicine* 1996; **154**:S234–S239.
3. Stacy EW. A generalization of the gamma distribution. *Annals of Mathematical Statistics* 1962; **33**:1187–1192.
4. Stacy EW, Mihram GA. Parameter estimation for a generalized gamma distribution. *Technometrics* 1965; **7**:349–358.
5. Prentice RL. A log gamma model and its maximum likelihood estimation. *Biometrika* 1974; **61**:539–544.
6. Holtz TH, Sternberg M, Kammerer S, Laserson KF, Riekstina V, Zarovska E, Skripconoka V, Wells CD, Lelmane V. Time to sputum culture conversion in multidrug-resistant tuberculosis: predictors and relationship to treatment outcome. *Annals of Internal Medicine* 2006; **144**:650–659.
7. Schneider MF, Gange SJ, Williams CM, Anastos K, Greenblatt RM, Kingsley L, Detels R, Muoz A. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984–2004. *AIDS* 2005; **19**:2009–2018.
8. Glaser RE. Bathtub and related failure rate characterizations. *Journal of the American Statistical Association* 1980; **75**:667–672.
9. Cox C, Ma G. Asymptotic confidence bands for generalized nonlinear regression models. *Biometrics* 1995; **51**:142–150.
10. DHHS/Henry J. Kaiser family foundation panel on clinical practices for the treatment of HIV infection. Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. May 4, 2006 version. Available at http://hivatis.org
11. Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo Jr CR. The multicenter AIDS cohort study: rationale, organization, and selected characteristics of the participants. *American Journal of Epidemiology* 1987; **126**:310–318.
12. Barkan SE, Melnick SL, Preston-Martin S, Weber K, Kalish LA, Miotti P, WIHS Collaborative Study Group. The women's interagency HIV study. *Epidemiology* 1998; **9**:117–1257.
13. Centers for disease control and prevention. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR* 1992; 1–19.
14. Smith RM, Bain LJ. An exponential power life-testing distribution. *Communications in Statistics—Theory and Methods* 1975; **4**:469–481.
15. Mudholkar GS, Srivastava DK, Kollia GD. A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association* 1996; **91**:1575–1583.
16. Prentice RL. Discrimination among some parametric models. *Biometrika* 1975; **62**:607–614.
17. Peng Y, Dear KBG, Denham JW. A generalized F mixture model for cure rate estimation. *Statistics in Medicine* 1998; **17**:813–830.
18. Ciampi A, Hogg SA, Kates L. Regression analysis of censored survival data with the generalized F family—an alternative to the proportional hazards model. *Statistics in Medicine* 1986; **5**:85–96.
19. Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modeling and estimation of treatment effects. *Statistics in Medicine* 2002; **21**:2175–2197.
20. Egger M, May M, Chene G, Phillips AN, Ledergerber B, Dabis F, Costagliola D, D'Arminio Monforte A, de Wolf F, Reiss P, Lundgren JD, Justice AC, Staszewski S, Leport C, Hogg RS, Sabin CA, Gill MJ, Salzberger B, Sterne JA. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet* 2002; **360**:119–129 (Erratum in: *Lancet* 2002; **360**:1178).
21. Putter H, Sasako M, Hartgrink HH, van de Velde CJH, van Houwelingen JC. Long-term survival with non-proportional hazards: results from the Dutch Gastric Cancer Trial. *Statistics in Medicine* 2005; **24**:2807–2821.
22. Klein JP, Moeschberger ML. *Survival Analysis*: *Techniques for Censored and Truncated Data* (2nd edn). Springer: New York, 2003.
23. Cleves MA, Gould WW, Gutierrez RG. *An Introduction to Survival Analysis Using Stata* (Revised edn). Stata Press: College Station, 2004.
24. Allison PA. *Survival Analysis Using SAS*: *A Practical Guide*. SAS Institute: Cary, NC, 1995.
25. Venables WN, Ripley BD. *Modern Applied Statistics with S* (4th edn). Springer: New York, 2002.
26. Gould W, Pitblado J, Sribney W. *Maximum Likelihood Estimation with Stata* (3rd edn). Stata Press: College Station, 2006.

DOI: 10.1002/sim