STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:4301–4312
Published online 14 April 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3292

The generalized F distribution: An umbrella for parametric survival analysis

Christopher Cox*, †

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room E7642, Baltimore, MD 21205, U.S.A.

SUMMARY

In a recent tutorial my colleagues and I advocated the generalized gamma (GG) distribution as a platform for parametric survival analysis. The GG family includes all four of the common types of hazard functions, making it particularly useful for estimating individual hazard functions as well as both relative hazards and relative times. In addition, the GG includes most of the commonly used parametric survival distributions. Survival analysis based on the GG distribution is practical since regression models are available in commonly used statistical packages. It is well known that the GG is contained in an even larger family, the generalized F (GF) distribution, which also includes the log logistic. The GF thus provides additional flexibility for parametric modeling. In this paper we discuss the GF family from this perspective. We provide a characterization of the hazard functions of the GF, showing that, except for the GG, the available hazard functions are limited to decreasing and arc-shaped hazards and, in particular, that the hazard function can be decreasing but not monotone. We also discuss fitting the GF with an alternative parameterization using standard statistical software and refine a description of the hazard functions for death after a diagnosis of clinical AIDS in four different eras of HIV therapy. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: survival analysis; parametric models; generalized gamma distribution; generalized F distribution; hazard function

*Correspondence to: Christopher Cox, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room E7642, Baltimore, MD 21205, U.S.A.
†E-mail: ccox@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 29 August 2007
Accepted 3 March 2008

1. INTRODUCTION

In a recent tutorial on parametric models for the analysis of right-censored and left-truncated survival data, Cox *et al.* [1] advocated the generalized gamma (GG) distribution as a standard platform for model fitting. This proposal was based on a taxonomy for the GG hazard functions, showing that the family includes all four of the most common shapes: monotonically increasing and decreasing as well as bathtub and arc-shaped hazards. Furthermore, the family includes most of the commonly used distributions: Weibull, log normal and gamma. Equally important, GG regression models are available in standard statistical packages such as SAS, Stata and S-Plus, making their application routine.

The use of the GG family offers the possibility of an improved fit compared with standard Weibull or log-normal models as well as a means of testing the goodness of fit of these simpler models. One approach to examining the goodness of fit of the GG itself is to graphically compare the estimated survival function with the product limit estimate, or the cumulative hazard with the Nelson–Aalen estimate [1]. Alternatively, the GG is a member of the even larger generalized F (GF) family, which also includes the commonly used log-logistic distribution. The use of the GF provides a way to assess the fit of a GG model, and the family has in fact been proposed for the parametric regression analysis of survival data [2]. If the family is to be used in this manner, then it is useful to characterize the behavior of the hazard functions as has already been done for the GG.

In this paper we provide a characterization of the hazard functions of the GF distribution. Somewhat disappointingly, except for the special case of the GG, the GF hazard can only be decreasing or arc-shaped, the same hazard behavior exhibited by the log logistic. We also show by example that the hazard can be decreasing with a relative minimum followed by a relative maximum before approaching a limit of zero as time becomes infinite, so that the function is not monotone decreasing. Thus, except for the limiting case of the GG sub-family, the hazard behavior of the GF distribution is somewhat limited. In spite of this limitation, the family does offer additional flexibility that may allow an improved fit, as our application shows.

A second limitation of the GF is that it is not presently available in the standard software packages for survival analysis. We discuss model fitting using general-purpose programs for maximum likelihood estimation. As a part of this discussion, we argue that a parameterization proposed by Prentice provides a natural linkage to the GG, is more useful for model fitting and is not more difficult to program. Finally, we illustrate the use of the GF by revisiting the application discussed in our original tutorial. This involved a comparison of survival after a clinical diagnosis of AIDS in four different eras of antiretroviral therapy at the population level. In one of the four therapy eras, the GF provided a significantly better fit, and we obtained an improved description of the hazard function, illustrating the utility of the GF as an umbrella over the GG family.

2. THE GENERALIZED F FAMILY

2.1. The GF distribution and its hazard functions

The GF distribution is a four-parameter family, generalizing the central F with noninteger degrees of freedom $(2m_1, 2m_2)$ by adding location (β) and scale $(\sigma>0)$ parameters according to the form of the standard accelerated failure time model [2]. Let $w(t)=[\log(t)-\beta]/\sigma$ so that $e^w=(e^{-\beta}t)^{1/\sigma}$ and $t=e^{\beta+\sigma w}$. Suppressing the dependence on the parameters $\theta=(\beta,\sigma,m_1,m_2)$, the density function

DOI: 10.1002/sim

becomes

$$f_{\mathrm{GF}}(t)=\frac{\mathrm{e}^{-\beta m_1/\sigma}t^{(m_1/\sigma)-1}(m_1/m_2)^{m_1}}{\sigma B(m_1,m_2)[1+(m_1/m_2)(\mathrm{e}^{-\beta}t)^{1/\sigma}]^{(m_1+m_2)}}=\frac{C(\theta)t^{(m_1/\sigma)-1}}{[1+(m_1/m_2)(\mathrm{e}^{-\beta}t)^{1/\sigma}]^{(m_1+m_2)}}$$

$$=\frac{1}{B(m_1,m_2)}\left(\frac{\frac{m_1}{m_2}\mathrm{e}^{w}}{1+\frac{m_1}{m_2}\mathrm{e}^{w}}\right)^{m_1}\left(\frac{1}{1+\frac{m_1}{m_2}\mathrm{e}^{w}}\right)^{m_2}\frac{1}{\sigma t}=\frac{1}{B(m_1,m_2)}P^{m_1}(1-P)^{m_2}\frac{1}{\sigma t}\quad(1)$$

where $P(t)=(m_1\mathrm{e}^{w}/m_2)/(1+m_1\mathrm{e}^{w}/m_2)=1/(1+m_2\mathrm{e}^{-w}/m_1)>0$, a strictly increasing function with $\lim_{t\to 0^+}P(t)=0$ and $\lim_{t\to\infty}P(t)=1$, and $B(m_1,m_2)$ is the beta function evaluated at $m_1,m_2>0$. The family includes the generalized log-logistic distribution ($m_1=m_2=m$), and the log logistic is the special case $m=1$. Each of the limiting cases $m_1\to\infty$ and $m_2\to\infty$ yields members of the GG family, and the combination produces the log normal, although with this parameterization the limiting distribution is degenerate. The GF also includes the Burr XII ($m_1=1$), used in econometrics and Burr III ($m_2=1$) families. Ciampi *et al.* [2] and Peng *et al.* [3] provide tree diagrams illustrating the various special cases of the GF distribution. If $T\sim\mathrm{GF}(\beta,\sigma,m_1,m_2)$ and $b>0$, then $\mathrm{e}^{a}T^{b}\sim\mathrm{GF}(a+\beta b,\sigma b,m_1,m_2)$; and if $b<0$ then $\mathrm{e}^{a}T^{b}\sim\mathrm{GF}(a+\beta b,\sigma|b|,m_2,m_1)$; in particular for $T\sim\mathrm{GF}(\beta,\sigma,m_1,m_2)$, we have $1/T\sim\mathrm{GF}(-\beta,\sigma,m_2,m_1)$; a similar relationship holds for the GG distribution [1].

Given the fact that the GF family contains the GG, one might hope that the richness of the hazard behavior of the GF for finite values of m_1 and m_2 would equal or surpass that of the GG. As shown in Appendix A and summarized in Table I, the shape of the GF hazard for the case $0<m_1,m_2<\infty$ is basically the same as the log logistic. Case (1) includes the log logistic with $1>\sigma$, and cases (2) and (3) cover the log logistic with $1=\sigma$ and $1<\sigma$, respectively. In particular, the shape of the GF hazard is determined by the parameters $m_1>0$ and $\sigma>0$, independently of m_2. Similar results were obtained by Marshall and Olkin [4] using an alternative parameterization and different methods. For $m_1<\sigma<1$, Figure 1 shows four different hazard functions for the generalized log-logistic distribution ($m_1=m_2=m$) with $\beta=0$ and $m\leqslant\sigma=0.50$, one of which, $\theta=(0,0.50,0.49,0.49)$, is not monotone decreasing. Note that the hazard for $m=\sigma=0.50$ is arc-shaped and satisfies $0<h(0)<\infty$ as required by case (2) in Table I. We conclude that the GG sub-family of the GF has a diversity of hazard shapes that is not matched by the remaining distributions in the family ($0<m_1,m_2<\infty$). In spite of this limitation, the GF does offer additional richness of distribution shapes and, in particular, the ability to check the goodness of fit of the GG and to compare it with the log logistic.

Table I. Hazard behavior of the GF distribution for finite values of the parameters (σ,m_1,m_2).

Parameters		Hazard function
(1) $m_1>\sigma$		Arc-shaped, $h(0)=0=h(\infty)$
(2) $m_1=\sigma$,	$\sigma\geqslant 1$	Decreasing, $h(0)=C(\theta)>0$, $h(\infty)=0$
	$\sigma<1$	Arc-shaped, $h(0)=C(\theta)>0$, $h(\infty)=0$
(3) $m_1<\sigma$,	$\sigma\geqslant 1$	Decreasing, $h(0^+)=\infty$, $h(\infty)=0$
	$m_1<\sigma<1$	Decreasing, not necessarily monotone, $h(0^+)=\infty$, $h(\infty)=0$

DOI: 10.1002/sim

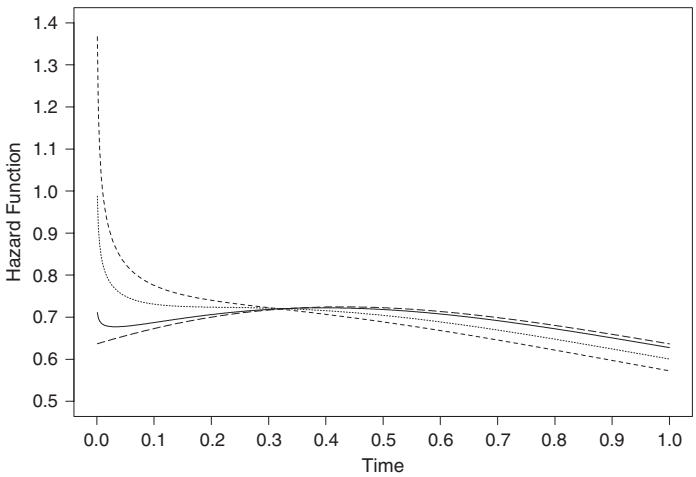

Figure 1. Four GF hazard functions for $\beta=0$, $\sigma=0.50$ in (1) and $m_1=m_2=m$: (top to bottom) $m=0.43$ (- - -), $m=0.46$ (.....), $m=0.49$ (——) and $m=0.50$ (– – –).

2.2. *Linking to the GG through an alternative parameterization*

Prentice proposed an alternative parameterization for the GF family [5] in order to produce a well-behaved likelihood for the limiting case of the GG, including the log normal, in the sense of having finite, nonzero derivatives of the log likelihood on the boundary, $m_1=\infty$ or $m_2=\infty$. Prentice also proposed what has become the standard parameterization for the GG for similar reasons [6], and the fact that this parameterization has been used in standard survival software illustrates its usefulness for model fitting. This parameterization replaces $0<m_1, m_2<\infty$ by alternative parameters

$$q=\left(\frac{1}{m_1}-\frac{1}{m_2}\right)\left(\frac{1}{m_1}+\frac{1}{m_2}\right)^{-1/2}, \quad p=\frac{2}{m_1+m_2}$$

Thus, we have $-\infty<q<\infty$ and $p>0$, and these parameters determine m_1 and m_2

$$m_1=2[q^2+2p+q(q^2+2p)^{1/2}]^{-1}, \quad m_2=2[q^2+2p-q(q^2+2p)^{1/2}]^{-1}$$

Note that with this definition, $q\geqslant 0$ is equivalent to $m_2\geqslant m_1$, and for $q<0$ we have $m_1>m_2$. To complete the parameterization, define $\delta=(m_1^{-1}+m_2^{-1})^{1/2}=(q^2+2p)^{1/2}$, and replace $\sigma\equiv\sigma/\delta$,

$$f_{\mathrm{GF}}(t)=\frac{\delta \mathrm{e}^{-\beta m_1/\sigma} t^{(\delta/\sigma)m_1}(m_1/m_2)^{m_1}}{t\sigma B(m_1,m_2)[1+(m_1/m_2)(\mathrm{e}^{-\beta}t)^{\delta/\sigma}]^{(m_1+m_2)}} \tag{2}$$

With this alternative parameterization the GG family is given by the limiting case $p=0$ [5], with $q=\lambda$, the shape parameter of the GG and $\delta=|\lambda|$. Specifically, if $p=0$ and $q>0$ then $m_2=\infty$, and consequently $q=1/m_1^{1/2}$ and $q=\delta=\lambda>0$ in the GG density

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

DOI: 10.1002/sim

Similarly, if $p=0$ and $q<0$ then $m_1=\infty$ and $q=-1/m_2^{1/2}$ so that $q=\lambda<0$ and $\delta=|\lambda|$. The limiting case $p=q=0$ is the log normal. The case $q=0$ defines the generalized log-logistic distribution with $m=1/p$, and $p=1$ is the special case of the log logistic. Similarly, the Burr XII distribution ($m_1=1$) is defined by $q=q(p)=(1-p)[2/(2-p)]^{1/2}$ ($p<2$), and $q=-q(p)$ defines the Burr III family ($m_2=1$). Both of these distributions include the Weibull ($p=0, q=1$) as well as the log logistic. The final substitution $\sigma\equiv\sigma/\delta$ is not necessary but has the advantage of ensuring that both GF and GG are estimating the same scale parameter, which is useful for model fitting and interpretation. On the other hand, with this parameterization the shape of the GF hazard depends on the ratio σ/δ, rather than σ.

The alternative parameterization is also consistent with that for the GG, for example, from case (1) of Table I, the GF hazard eventually becomes arc-shaped for $q\leqslant 0$ as $m_1\to\infty$, consistent with the arc-shaped GG hazard for $\lambda\leqslant 0$. For $q>0$, the most interesting case is where the limiting GG hazard has a bathtub shape, $\lambda>\max\{\sigma, 1/\sigma\}$. Figure 2 shows a series of GF hazard functions for decreasing values of p corresponding to increasing values of m_2, so that the limiting distribution is GG with $\lambda>0$. For all of these examples, $m_1=0.49<0.50=\sigma/\delta<1$, so that the hazard is decreasing, although not necessarily monotone. The hazard for $m_2=8000$ is indistinguishable from that of the GG with $\lambda=1/0.49^{1/2}=1.4286=\delta$ and $\sigma=0.50\delta=0.7143$. Since $\lambda>1/\sigma=1.4$, the limiting GG distribution has a bathtub-shaped hazard function.

These results can be represented graphically using a schematic that is similar to the hazard taxonomy for the GG presented by Cox *et al.* [1], providing a representation of the shapes of the GF hazard in a neighborhood of the limiting GG, i.e. as $p\to 0^+$. The simplest way to accomplish this is in terms of the (original) parameters $\sigma_F=\sigma/\delta$ and m_1, since these determine the shape of the hazard. For the case $q>0$, the limiting GG distribution has $\sigma=\sigma_F\delta=\sigma_F/m_1^{1/2}$ and $\lambda=1/m_1^{1/2}$. We can use this property to map the GF parameters (σ_F, m_1) to the GG parameters (σ, λ). For example, the curve defined by $\lambda=1/\sigma$, which defines the *ammag* distribution and is an essential element of the GG hazard taxonomy, becomes $m_1=\sigma_F(m_1\delta=\sigma)$. The generalized *ammag*, defined

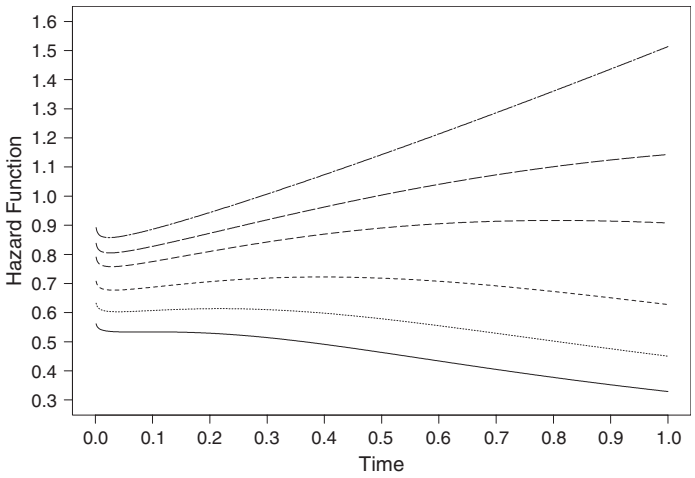

Figure 2. GF hazard functions for $\beta=0$, $\sigma/\delta=0.50$ in (2) and $m_1=0.49$: (bottom to top) $m_2=0.2$ (—), $m_2=0.3$ (.....), $m_2=0.49$ (_ _ _ _), $m_2=1$ (- - - -), $m_2=2$ (– – –), $m_2=8000\approx \mathrm{GG}(0,\sigma,1/m_1^{1/2})$ (–·–·–).

DOI: 10.1002/sim

by this condition for $p>0$, is the only member of the GF family whose density and hazard functions satisfy $0<h(0)=C(\theta)<\infty$ (see Appendix A). Similarly, the curve $\lambda=\sigma$, which defines the gamma distribution, becomes $\sigma_F=1$ ($\sigma=\delta$), for $p>0$ the special case of the F distribution. For the case $q\leqslant 0$, as m_1 increases the GF hazard is eventually arc-shaped, corresponding to the limiting GG with $\lambda\leqslant 0$.

Figure 3 contains the full representation and should be compared with Figure 1 of Cox *et al.* [1], which presents the hazard taxonomy for the GG. For $p\approx 0$, the curve $m_1=\sigma_F$ divides the (σ,λ) half-plane into two regions; in the upper region the GF hazard is decreasing and in the lower region the hazard is arc-shaped. In the limit, the GG hazard has a different shape in each of the four regions determined by $\lambda=1/\sigma$ ($m_1=\sigma_F$) and $\lambda=\sigma$ ($\sigma_F=1$). Below the line $\sigma_F=1$, the shape of the hazard is the same as the GF; above this line the GG hazard has a bathtub shape in the upper region and is increasing in the left-hand region of the graph. A nearby GF hazard will be decreasing but not monotone in the upper region. The various special cases of the GG distribution are also included in Figure 3.

2.3. Analysis of survival data using the GF family

As indicated previously, the GF distribution is not available in any of the standard statistical packages for parametric survival analysis. We therefore consider fitting the GF using general purpose computer programs for maximum likelihood estimation. Because of its regularity properties and consistency with the GG, we will use parameterization (2). The corresponding survival function

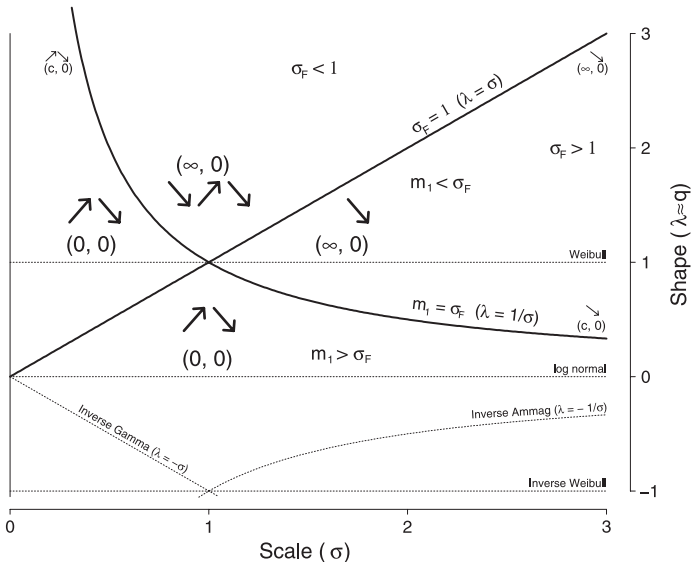

Figure 3. Shapes of the GF hazard for $p\approx 0$ (large values of m_2 corresponding to $q>0$ and large m_1 for $q<0$). The shape of the hazard function is determined by the parameters m_1 and σ_F (Table I). For $q>0$, we have $q\approx 1/m_1^{1/2}=\lambda$ and $\sigma_F\approx\sigma/\lambda$. Alternatively, for $q<0$, $q\approx -1/m_2^{1/2}=\lambda$ and $\sigma_F\approx\sigma/|\lambda|$. Above the curve $m_1=\sigma_F$ corresponding to $\sigma=1/\lambda$, the hazard is decreasing; below this curve it has an arc shape. Below the line $\sigma_F=1$ corresponding to $\sigma=\lambda$ in the upper region, the hazard is monotone. The figure also includes various special cases of the limiting ($p=0$) GG family.

DOI: 10.1002/sim

can be expressed in terms of the cumulative distribution function (CDF) of the beta distribution [3] and the functions $m_1(p,q)$ and $m_2(p,q)$

$$S_{\mathrm{GF}}(t)=\int_0^{m_2(m_2+m_1e^w)^{-1}} x^{m_2-1}(1-x)^{m_1-1}/B(m_2,m_1)\,\mathrm{d}x \tag{3}$$

Expressions (2) and (3) can be used to compute the log likelihood for left-truncated and right-censored observations using supplied functions for the beta density and CDF.

For SAS PROC NLMIXED sample expressions are as follows. The `parms` statement defines the four parameters and provides initial values, and `ll` is the log likelihood for failures (`event = 1`) and right-censored observations (`event = 0`). The variable lot is the log of the failure or censoring time.

```
proc nlmixed cov fd;
 parms beta=0 sigma=1 q=1 p=1; * p=.000001; * q = 0;
 m1 = 2/(q*q + 2*p + q*sqrt(q*q + 2*p));
 m2 = 2/(q*q + 2*p - q*sqrt(q*q + 2*p));
 d = sqrt(q*q + 2*p);

 if (event = 1) then
    ll = m1*d*(lot - beta)/sigma + m1*log(m1/m2)
         - log(sigma/d) - log(beta(m1,m2))
         - (m1 + m2)*log(1 + m1*exp(d*(lot-beta)/sigma)/m2);

 if (event = 0) then
   ll = log( cdf('beta',m2/(m2 + m1*exp(d*(lot-beta)/sigma)),m2,m1) );
model lot ~ general(ll); run;
```

The GG distribution can be approximated by removing `p` from the parameter list and setting its value to a small positive number. Removing `q` and setting $q=0$ give the generalized log-logistic distribution, and removing both `p` and `q` and setting `m1` and `m2` both equal to one (corresponding to $q=0$ and $p=1$) give the log-logistic distribution, although to obtain the standard parameterization one must also set `d = 1;` instead of `d = sqrt(2);`. The use of the SAS NLMIXED procedure is discussed by Cox *et al.* [1] and additional details can be found there.

3. AN APPLICATION: CHECKING THE FIT OF THE GG

We revisit the application discussed by Cox *et al.* [1]. This involved the analysis of time from a diagnosis of clinical AIDS to death in four different eras of HIV therapy. 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 known as HAART (DHHS/Henry J. Kaiser Family Foundation Panel, 2006). Since the evolution of HIV therapy followed a definite pattern over time, it is possible to define sequential calendar periods corresponding to distinct therapeutic eras. Cox *et al.* [1] distinguished four such periods, the first 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

DOI: 10.1002/sim

the (short- to moderate-term stable) era of HAART (July 1998–December 2003). This approach allowed assessment of the changing pattern of survival at the population level after the development of clinical AIDS from 1984 through the HAART eras. For further discussion of these data, see [1].

In these analyses, goodness of fit of the GG distribution was evaluated by comparison of the log of the cumulative hazard with that based on the Kaplan–Meier estimate. The fit appeared adequate, with the possible exception of the second period, where there was a small discrepancy between the GG and the nonparametric estimate in the right tail of the distribution, where the number of events was relatively small.

We re-examined the goodness of fit for the second period by fitting the GF distribution to the data. This period had 660 subjects, with 445 events (67 per cent). The value of -2 times the log likelihood for the four-parameter GF was 1597.6, compared with 1604.3 for the three-parameter GG, corresponding to a likelihood ratio statistic of $\chi^2_1 = 6.72$ ($p = 0.0095$), indicating that the GF provided a significantly better fit. The values of the parameters (SE) were $\beta = 0.5002$ (0.0755), $\sigma = 0.6387$ (0.1006), $q = 0.4538$ (0.1939) and $p = 2.2938$ (1.2587). The NLMIXED procedure also has the ability to estimate nonlinear functions of the parameters, with standard errors computed by the delta method. Using this feature, an asymmetric 95 per cent confidence interval for p based on the log transformation was $(0.78, 6.73)$. The degrees of freedom were $m_1 = 0.3456$ and $m_2 = 0.5263$, so that $\delta = 2.1894$ and $\sigma_F = \sigma/\delta = 0.2917 < m_1$, and the hazard function is arc-shaped. The GG parameter estimates obtained by Cox *et al.* [1] using a similar approach based on the gamma density and CDF were $\beta = 0.6493$ (0.0603), $\sigma = 0.8475$ (0.0485) and $\lambda = 0.8510$ (0.1267); as one might expect, the final model for this period was a gamma distribution ($\lambda = \sigma$), with an increasing hazard. When the parameter p was removed from the parameter list of the GF program and set to 0.000001, the results were nearly identical to those for the GG, $\beta = 0.6492$ (0.0606), $\sigma = 0.8475$ (0.0488) and $q = \lambda = 0.8509$ (0.1277), despite the fact that the value of m_2 used with the supplied beta function was very large. The value of -2 times the log likelihood for the generalized log-logistic distribution was 1602.4, with a likelihood ratio statistic of $\chi^2_1 = 4.75$ ($p = 0.029$). The value for the Burr XII distribution was similar, 1601.9, whereas the Burr III distribution actually provided an adequate fit (1598.5), also with an arc-shaped hazard. For the log logistic the value was 1615.1, with a likelihood ratio statistic of $\chi^2_2 = 17.46$ ($p = 0.00016$), indicating that the log logistic did not provide a satisfactory fit.

Figure 4 shows the GF, GG and log-logistic hazard functions, together with a nonparametric estimate based on kernel smoothing, with pointwise 95 per cent confidence bands [7]. Consistent with the likelihood ratio test, this estimate and the confidence bands show that the GF does indeed provide a better description of the hazard function for these data; the confidence bands also highlight the fact that there are relatively few events after 3.5 years. The nonparametric estimate clearly indicates that the hazard function has an arc shape similar to that of the GF but not the GG. As noted by a reviewer, the graph also suggests that the fit of the GF may not be fully satisfactory, although the width of the confidence bands makes this somewhat difficult to assess.

4. DISCUSSION

We have discussed the GF distribution as an umbrella over both the GG and log-logistic distributions for the analysis of survival data. Although the hazard behavior of the GF for $p > 0$ is limited to arc-shaped or decreasing hazards, the family does offer additional flexibility that may be useful in practice, as indicated by the application. From a practical point of view, the behavior of the

DOI: 10.1002/sim

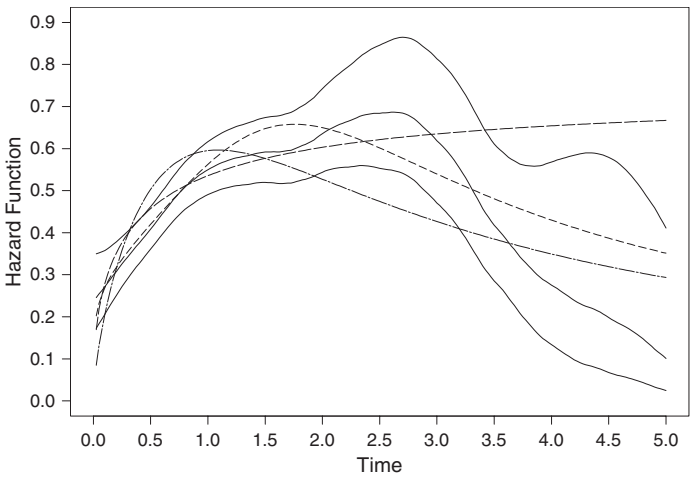

Figure 4. Hazard functions for the GF (- - -), GG (– – –) and log-logistic (–·–·–) distributions with a nonparametric estimate based on kernel smoothing, including 95 per cent confidence intervals (——).

hazard function is most relevant in the range of the data, and from this perspective the GF family can approximate all four of the common hazard types.

Fitting the GF distribution to left-truncated and right-censored survival data is relatively straightforward using programs such as NLMIXED and taking advantage of supplied functions for the beta density and CDF. Similar programs are available in S-Plus [8] and Stata [9]. The use of the parameterization proposed by Prentice [5] provides a smooth linkage with the GG family and also helps to avoid the kind of convergence problems noted by Ciampi *et al.* [2]. These same computational tools can be used to compute the survival and hazard functions for plotting as well as both (nonproportional) relative hazards and relative times, see [1]. Both the data and programs discussed in that paper are available on a web site at http://statepi.jhsph.edu/software; using the programming statements in Section 2.3, the latter can be easily adapted to fit GF survival models.

APPENDIX

We use the results of Glaser [10] to study the shape of the hazard function of the generalized *F* distribution for the case $p>0$, corresponding to $0<m_1, m_2<\infty$. Form (1) for the GF density is most convenient for this purpose. We begin by recording the limiting behavior of the hazard function

$$\begin{aligned}\varepsilon = \lim_{t\to 0^+} h(t) = \lim_{t\to 0^+} f(t) &= 0 && \text{if } m_1>\sigma\\ &= C(\theta) && \text{if } m_1=\sigma\\ &= \infty && \text{if } m_1<\sigma\end{aligned}$$

DOI: 10.1002/sim

Following Glaser [10], let $\eta(t) = -f'(t)/f(t)$; we have

$$\eta(t) = -\left(\frac{m_1}{\sigma} - 1\right) t^{-1} + (m_1 + m_2)[1 + (m_1/m_2)(\mathrm{e}^{-\mu}t)^{1/\sigma}]^{-1} m_1/(m_2\sigma)(\mathrm{e}^{-\mu}t)^{1/\sigma} t^{-1}$$
$$= -\left(\frac{m_1}{\sigma} - 1\right) t^{-1} + (m_1 + m_2) P(t)(\sigma t)^{-1} \tag{A1}$$

It follows that $\lim_{t\to\infty} h(t) = \lim_{t\to\infty} \eta(t) = 0$. Thus, at least in a limiting sense, the hazard function of the GF distribution for the case $0 < m_1, m_2 < \infty$ is decreasing for $m_1 \leqslant \sigma$, and arc-shaped (UBT in Glaser's notation) for $m_1 > \sigma$. A special case is the log-logistic distribution ($m_1 = m_2 = 1$), for which the hazard is decreasing for $1 \leqslant \sigma$, and arc-shaped for $1 > \sigma$.

We now turn to the behavior of the hazard function between the limiting values. To apply the results of Glaser, we must examine the behavior of $\eta'(t)$. Using the fact that $P'(t) = P(t)[1 - P(t)]/(\sigma t)$, we have

$$\eta'(t) = \frac{1}{\sigma^2 t^2}[-(m_1 + m_2) P(t)^2 + (1 - \sigma)(m_1 + m_2) P(t) + \sigma(m_1 - \sigma)] = \frac{\phi(P)}{\sigma^2 t^2}$$

The expression in square brackets, $\phi(P)$, is a quadratic polynomial in $P(t)$ having a maximum at $P = (1 - \sigma)/2$, and at most two roots in the interval $0 < P < 1$. Furthermore,

$$\lim_{t\to 0^+} \eta'(t) = \infty \qquad \text{if } m_1 > \sigma$$
$$= -\infty \quad \text{if } m_1 < \sigma \tag{A2}$$

The case $m_1 = \sigma$ is more complicated. For $c = (m_1 + m_2)/\sigma^2$ and $w = [\log(t) - \beta]/\sigma$,

$$\lim_{t\to 0^+} \eta'(t) = c \lim_{t\to 0^+} \frac{1}{t^2}\{-P(t)^2 + (1 - \sigma) P(t)\}$$
$$= -c \lim_{w\to -\infty} \left\{ \frac{e^{-2(\mu+\sigma w)}}{\left(1 + \dfrac{m_2}{\sigma}\mathrm{e}^{-w}\right)^2} - \frac{(1 - \sigma)\mathrm{e}^{-2(\mu+\sigma w)}}{1 + \dfrac{m_2}{\sigma}\mathrm{e}^{-w}} \right\}$$
$$= -c \lim_{w\to -\infty} \frac{1}{\left(\mathrm{e}^{\mu+\sigma w} + \dfrac{m_2}{\sigma}\mathrm{e}^{\mu+(\sigma-1)w}\right)^2} + c \lim_{w\to -\infty} \frac{1 - \sigma}{\mathrm{e}^{2(\mu+\sigma w)} + \dfrac{m_2}{\sigma}\mathrm{e}^{2\mu+(2\sigma-1)w}}$$
$$= 0^+ \qquad \text{if } 0 < \sigma < 1/2$$
$$> 0 \qquad \text{if } \sigma = 1/2$$
$$= \infty \qquad \text{if } 1/2 < \sigma < 1$$
$$< 0 \qquad \text{if } \sigma = 1$$
$$= -\infty \quad \text{if } \sigma > 1 \tag{A3}$$

As with the hazard it is clear that $\lim_{t\to\infty} \eta'(t) = 0$. Since $\lim_{P\to 1} \phi(P) = -\sigma(m_2 + \sigma) < 0$, there must exist $t_\theta > 0$ such that for $t > t_\theta$, we have $\eta'(t) < 0$. Using (2.4) in Glaser we have that $h(t)$ must eventually be strictly decreasing ($h'(t) < 0$ for $t > t_\theta$).

We must now consider three special cases. Suppose first that $m_1 > \sigma$. In this case, from (A2), $\phi(P)$ has exactly one root in the interval $0 < P < 1$, so that there must exit $t_0 > 0$ such that $\eta'(t) > 0$

DOI: 10.1002/sim

for $t<t_0$, $\eta'(t_0)=0$ and $\eta'(t)<0$ for $t>t_0$. This is condition (2.6) of Glaser. Furthermore, from the fact that in this case $\varepsilon=0$, we conclude from lemma a(iii) that the hazard is arc-shaped. The log-logistic distribution with $1>\sigma$ is a special case.

For the case $m_1=\sigma$, $\phi(P)$ has a root at $P=1-\sigma$. Thus, if $\sigma\geqslant 1$ we have from (A3) that $\eta'(t)<0$ for all $t>0$. From Theorem (b) of Glaser it follows that the hazard is decreasing (D). This case includes the log-logistic distribution with $\sigma=1$. If $\sigma<1$ then $\phi(P)$ has exactly one root and condition (2.6) of Glaser is again satisfied. As noted by Glaser the behavior of the hazard in this case is determined by the behavior of the function $g(t)=1/h(t)$ near the origin, specifically whether $g(t)$ is initially increasing or decreasing. We have that $g'(t)=g(t)\eta(t)-1$, and we have already observed that in this case

$$\lim_{t\to 0^+} g(t)=1/C(\theta)>0$$

Thus, we only need

$$\lim_{t\to 0^+}\eta(t)=\lim_{t\to 0^+}(\sigma+m_2)\frac{P(t)}{\sigma t}=\lim_{w\to-\infty}\frac{(\sigma+m_2)}{\sigma}\frac{1}{e^{\mu+\sigma w}+\dfrac{m_2}{\sigma}e^{\mu+(\sigma-1)w}}=0$$

Thus $g(t)$ is initially decreasing, and from lemma b(iii) of Glaser it follows that $h(t)$ is arc-shaped. In summary, the case $m_1=\sigma$ includes both possible shapes for the hazard function.

Finally, if $m_1<\sigma$, then again from (A2), $\phi(P)$ either has no roots or two roots in the interval $0<P<1$. If $\sigma\geqslant 1$, then two roots is impossible so that the answer is none; in this case we have from (A2) that $\eta'(t)<0$ for all $t>0$, and it follows that the hazard is decreasing (D). An example is the log-logistic distribution with $\sigma>1$. The remaining possibility is $m_1<\sigma<1$. If $\phi(P)$ has no roots, then the hazard is again decreasing by (A2). In order to have two roots the maximum of $\phi(P)$ must be positive or equivalently the discriminant in the quadratic formula must be positive, yielding $(1-\sigma)^2(m_1+m_2)>4\sigma(\sigma-m_1)$. In this case both roots are included in the interval $0<P<1$, and the results of Glaser cannot be applied. In the special case $m_1=m_2=m$, there are two roots provided $m>2\sigma^2/(1+\sigma^2)$, which is satisfied, for example, by $\sigma=0.50$ and $m=0.49>0.40$, shown in Figure 1.

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). I would like to thank Alvaro Muñoz for his imagination and insights into parametric survival models in particular and survival analysis in general.

Studies were approved by the Committees of Human Research of participating institutions. The MACS web site is located at http://statepi.jhsph.edu.

REFERENCES

1. Cox C, Chu H, Schneider MF, Muñoz A. Tutorial in biostatistics: parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine* 2007; **26**:4352–4374.
2. 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.

DOI: 10.1002/sim

3. Peng Y, Dear KBG, Denham JW. A generalized *F* mixture model for cure rate estimation. *Statistics in Medicine* 1998; **17**:813–830.
4. Marshall AW, Olkin I. *Life Distributions*. Springer: New York, 2007.
5. Prentice RL. Discrimination among some parametric models. *Biometrika* 1975; **62**:607–614.
6. Prentice RL. A log gamma model and its maximum likelihood estimation. *Biometrika* 1974; **61**:539–544.
7. Klein JP, Moeschberger ML. *Survival Analysis*: *Techniques for Censored and Truncated Data* (2nd edn). Springer: New York, 2003.
8. Venables WN, Ripley BD. *Modern Applied Statistics with S* (4th edn). Springer: New York, 2002.
9. Gould W, Pitblado J, Sribney W. *Maximum Likelihood Estimation with Stata* (3rd edn). Stata Press: College Station, 2006.
10. Glaser RE. Bathtub and related failure rate characterizations. *Journal of the American Statistical Association* 1980; **75**:667–672.

DOI: 10.1002/sim