Published in final edited form as:

Stat Med. 2014 September 20; 33(21): 3772–3780. doi:10.1002/sim.6159.

A Comparison of the Generalized Gamma and Exponentiated Weibull Distributions

Christopher Cox and **Matthew Matheson**

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland

SUMMARY

This paper provides a comparison of the three-parameter exponentiated Weibull (EW) and generalized gamma (GG) distributions. The connection between these two different families is that the hazard functions of both have the four standard shapes (increasing, decreasing, bathtub and arc-shaped), and in fact the shape of the hazard is the same for identical values of the three parameters. For a given EW distribution, we define a matching GG using simulation and also by matching the 5th, 50th and 95th percentiles. We compare EW and matching GG distributions graphically and using the Kullback-Leibler distance. We find that the survival functions for the EW and matching GG are graphically indistinguishable, and only the hazard functions can sometimes be seen to be slightly different. The Kullback-Leibler distances are very small and decrease with increasing sample size. We conclude that the similarity between the two distributions is striking, and therefore that the EW represents a convenient alternative to the GG with the identical richness of hazard behavior. More importantly, these results suggest that having the four basic hazard shapes may to some extent be an important structural characteristic of any family of distributions.

Keywords

Parametric survival; generalized gamma distribution; exponentiated Weibull distribution

1. INTRODUCTION

In a recent tutorial [1] we advocated the generalized gamma (GG) distribution as a platform for parametric survival analysis, with the resulting description of the effects of exposures by possibly non-proportional relative times (quantiles) as well as relative hazards, which can also be non-proportional. In this context, which may not be so uncommon in practice, richness of hazard shapes is an important characteristic. One of the most appealing features of the GG family is that it contains all four of the basic hazard shapes: increasing (I), decreasing (D), bathtub (B) and arc-shaped (A) hazard functions. In addition the GG distribution is available in standard statistical software packages, with various combinations of censoring and late entry. For further discussion see [1].

In this paper we consider another three-parameter family that also includes the four basic hazard shapes, the exponentiated Weibull distribution [2]. Although the EW distribution is not available in standard statistical packages for parametric survival analysis, methods discussed in [1] can be used to fit this parametric family. The goal of this paper is to compare the two families in order to determine whether additional flexibility in the shape of the hazard might be offered by considering both families of distributions. Our comparisons include fitting the two distributions to data on survival after an initial diagnosis of AIDS, to see whether more subtle aspects of the shape of the hazard function might be revealed.

2. TWO FAMILIES OF PARAMETRIC SURVIVAL DISTRIBUTIONS

2.1 The generalized gamma distribution

The GG is a three-parameter $(\beta, \sigma > 0, \kappa)$ family whose survival function is as follows.

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

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

where $\Gamma(t; \gamma) = \int_0^t x^{\gamma-1} e^{-x} dx / \Gamma(\gamma)$ is the cumulative distribution function for the gamma distribution with mean and variance equal to $\gamma > 0$. The corresponding density function is

$$f_{GG}(t) = \frac{|\kappa|}{\sigma t \Gamma(\kappa^{-2})} \left\{ \kappa^{-2} (e^{-\beta t})^{\kappa/\sigma} \right\}^{\kappa^{-2}} \exp \left\{ -\kappa^2 (e^{-\beta t})^{\kappa/\sigma} \right\}$$

The quantile function, which is useful for studying the two families of distributions, is

$$\begin{aligned} \log [t_{GG(\beta, \sigma, \kappa)}(p)] &= \beta + \sigma \log [t_{GG(0.1, \kappa)}(p)] = \beta + \frac{\sigma}{\kappa} \log [\kappa^2 \Gamma^{-1}(p; \kappa^{-2})] & \text{if } \kappa > 0 \\ \log [t_{GG(\beta, \sigma, \kappa)}(p)] &= \beta + \sigma \log [t_{GG(0.1, \kappa)}(p)] = \beta + \frac{\sigma}{\kappa} \log [\kappa^2 \Gamma^{-1}(1-p; \kappa^{-2})] & \text{if } \kappa < 0 \end{aligned} \quad (1)$$

As discussed in [1] the GG family offers a useful platform for the parametric analysis of survival data. First, it is available in standard statistical packages such as SAS, Stata, S-Plus and R. Second, the GG includes as special cases the Weibull ($\kappa = 1$), lognormal ($\kappa = 0$), and gamma ($\kappa = \sigma$) distributions. Third, the hazard functions of the GG family include all four of the basic shapes.

- (1) increasing (I) hazard for $0 < \sigma < 1$ and $\sigma < \kappa/1/\sigma$
- (2) decreasing (D) hazard for $\sigma > 1$ and $1/\kappa < \sigma$
- (3) bathtub (B) hazard for $\kappa > \max(\sigma, 1/\sigma)$
- (4) arc-shaped (A) or unimodal hazard for $\kappa < \min(\sigma, 1/\sigma)$

For a graphical representation of these relationships see Figure 1 of [1].

2.2 The exponentiated Weibull distribution

This family was defined by Mudholkar and Srivastava [2] as the Weibull distribution with an additional resilience [3] parameter (given a distribution function F_1 , the family $F_\kappa = F_1^\kappa$), having three parameters ($\beta, \sigma > 0, \kappa > 0$) and the following survival function.

$$S_{EW}(t) = 1 - \left[1 - \exp \left\{ -\kappa^{-2} (e^{-\beta t})^{\kappa/\sigma} \right\} \right]^{1/\kappa^2}$$

We have again chosen the standard accelerated failure time parameterization, with slight modifications of the original notation (using the parameterization of Mudholkar and Srivastava, $\alpha = \kappa / \sigma$, $\theta = 1 / \kappa^2$; $\sigma = e^{\beta + 2\sigma \log(\kappa)/\kappa}$) to align the parameterizations of the two families. To compute the hazard we also need the density function.

$$f_{EW}(t) = \frac{1}{\kappa^3 \sigma t} \left[1 - \exp \left\{ -\kappa^{-2} (e^{-\beta t})^{\kappa/\sigma} \right\} \right]^{1/\kappa^2 - 1} \exp \left\{ -\kappa^{-2} (e^{-\beta t})^{\kappa/\sigma} \right\} (e^{-\beta t})^{\kappa/\sigma}$$

The EW quantile function is

$$\log [t_{EW(\beta, \sigma, \kappa)}(p)] = \beta + \sigma \log [t_{EW(0.1, \kappa)}(p)] = \beta + \frac{\sigma}{\kappa} \log \left\{ -\kappa^2 \log (1 - p^{\kappa^2}) \right\} \quad (2)$$

Mudholkar and Hutson [4] determined the hazard behavior of the EW family. With this parameterization, the hazard behavior of the EW distribution is also given by (1)-(4) above. The special case $\kappa = \sigma$ is the exponential distribution with a resilience parameter. The hazard function of the exponentiated exponential (EE) distribution has the same horizontal asymptote as the gamma distribution, $\lim_{t \rightarrow \infty} h(t) = \kappa^{-2} e^{-\beta}$ [3, Section 9.C]. For the EW family, the power parameter κ plays a similar role as the shape parameter of the GG family and will be referred to as the shape parameter for the purposes of our comparison. Note that for both three-parameter families the shape of the hazard is independent of the location parameter β . Thus we have two distinct families of distributions whose hazard functions include all four of the basic shapes, potentially offering a real increase in the variety of hazard functions for parametric survival models.

2.3 A comparison of the EW and GG distributions

Surprisingly, an initial graphical comparison of the results of fitting both distributions to the same data sets using the method of maximum likelihood as described in [1] showed that both the fitted survival curves and density functions were indistinguishable, and the hazard functions were only slightly different. Based on these initial results, we conducted a comparison of the two families of survival distributions from the following perspective: given an EW distribution, find a GG that closely approximates it, and compare the two distributions. In our first study, the parameters of the approximating GG distribution were determined by simulation. We generated 1000 samples from a given EW distribution and estimated the GG parameters for each sample by maximum likelihood. The averages of the 1000 estimates, which approximate the expected values of the maximum likelihood

estimates, were taken as the parameters of the approximating GG. A series of 8 different EW distributions was chosen, with $\beta = 0$ and values of σ and κ selected so that each of the four different hazard shapes would be represented twice. The values of these parameters are shown in the left margin of Table 1 and are plotted in Figure 1. We chose three different sample sizes, 200, 800 and 3200, corresponding to (relatively) small, medium and large.

It is well known that when the correct model is $g(t)$, the parameters estimated by fitting the wrong model $[f(t)]$ by maximum likelihood are those minimizing the Kullback-Leibler (K-L) distance [5], $D_g(f) = E_g \{ \log[g(T)/f(T)] \}$. We compared the EW with the approximating GG by computing $D_{EW(\beta, \sigma, \kappa)} \left[GG(\hat{\beta}, \hat{\sigma}, \hat{\kappa}) \right]$ using numerical integration, and by plotting the integrand of the K-L distance. From the expression for the ratio of the two densities it is easy to see that if we start with the special case $EW(0, 1, \kappa)$ and denote the corresponding GG parameters as $GG(\hat{\beta}_0, \hat{\sigma}_0, \hat{\kappa}_0)$, then for $EW(\beta, \sigma, \kappa)$, we will have $GG(\beta + \sigma\hat{\beta}_0 + \sigma\hat{\sigma}_0, \hat{\kappa}_0)$ and the value of $D_{EW(\beta, \sigma, \kappa)}(\hat{\beta}, \hat{\sigma}, \hat{\kappa})$ will be unchanged, so that the K-L distance depends only on κ .

As a second approach, we selected a matching GG distribution for a given $EW(\beta, \sigma, \kappa)$ distribution by equating the 5th, 50th and 95th percentiles. These values were chosen because they produced relatively (to other choices) good agreement in terms of the K-L distance between the matched GG and that determined by simulation. Using substitution, we obtain a single nonlinear equation in the GG shape parameter. Denoting the three EW log percentiles (2) by $w_{.05}$, $w_{.50}$, $w_{.95}$ we have the following equation.

$$(w_{.50} - w_{.05}) \left[\log \left\{ t_{GG(0.1, \kappa)}(0.95) \right\} - \log \left\{ t_{GG(0.1, \kappa)}(0.50) \right\} \right] \\ - (w_{.95} - w_{.50}) \left[\log \left\{ t_{GG(0.1, \kappa)}(0.50) \right\} - \log \left\{ t_{GG(0.1, \kappa)}(0.05) \right\} \right] = 0$$

It follows from the form of this equation that the value of the approximating GG shape parameter depends only on the assumed value of the EW shape parameter. Therefore for comparison of the two families we can again consider only the special case $EW(0, 1, \kappa)$. It is true, however, that the difference between the two log quantile functions, and hence the error in the approximation, although independent of β , is proportional to σ . To study the effectiveness of this matching algorithm, we computed the K-L distance between the EW distribution and the matched GG for the eight examples in our simulation study. We then computed the K-L distance for each pair of matched distributions across a range of values of the $EW(0, 1, \kappa)$ shape parameter.

To provide a comparison involving actual data, we turn to an 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 in the Multicenter AIDS Cohort Study (MACS). 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. distinguished four such periods; we will focus on the second period, involving mono or combination therapy (January 1990 – December 1994). Data are available from 660 MACS participants who had a diagnosis of clinical AIDS during this period, with 445 deaths (67%). For further discussion of these data see [1]. This period was chosen because the GG distribution displayed a slight lack of fit in the original analysis, although in fact a satisfactory fit was provided by a standard gamma ($\sigma = \kappa$) distribution.

Both the EW and GG distributions are members of a larger family, the four-parameter exponentiated generalized gamma (EGG), which adds the resilience parameter to the GG family [6]. The survival function may be written as

$$S_{GG}(t) = 1 - \Gamma \left[\kappa^{-2} \left(e^{-\beta t} \right)^{\kappa/\sigma}; \kappa^{-2} \right]^{1/\theta^2} \quad \text{if } \kappa > 0$$

$$S_{GG}(t) = 1 - \left\{ 1 - \Gamma \left[\kappa^{-2} \left(e^{-\beta t} \right)^{\kappa/\sigma}; \kappa^{-2} \right]^{1/\theta^2} \right\} \quad \text{if } \kappa < 0$$

The hazard behavior of this extended family clearly involves the scale, shape and power (resilience) parameters. Another way to compare the EW and GG distributions is to consider both as special cases of the EGG. This allows a comparison of the goodness-of-fit of each family to the same data set.

3. RESULTS OF THE COMPARISONS

3.1 Comparisons based on simulation and matching

The results of our simulation studies are summarized in the first three columns of Table 1. These include the means and standard deviations of the three GG parameters estimated from the simulated EW samples. The three sets of parameter estimates are similar, and the standard deviations decrease with increasing sample size. The table also contains the K-L distance from the GG distribution determined by these estimates to the true EW distribution. For comparison we also included the K-L distance from the corresponding fitted EW distribution to the true EW, also obtained by averaging fitted values for the simulated samples. The K-L distances of the fitted GG distributions are very small, and are clearly similar to the K-L distances of the fitted EW distributions. As one might expect, both sets of distances become smaller as the sample size increases, reflecting in part the decreasing bias of the maximum likelihood estimates. Note also that, as expected, the K-L distances for EW parameters with the same true value of κ are very similar. The simulation results are also represented graphically in Figure 1. In addition to the selected EW parameter values (σ, κ), the figure includes the corresponding values from Table 1 for the three different sample sizes, as well as the values from percentile matching. For comparison we also plotted the estimated EW parameters for the sample size of 200.

The final column of Table 1 contains the three GG parameters obtained by percentile matching, together with the corresponding K-L distance. The parameter estimates and the distances are close to those for the largest sample size (3200) from the simulation study; this can also be seen graphically in Figure 1. The figure also shows that there are clear

differences between the selected EW parameters and both the estimated and matched GG parameters, especially for the smallest value of κ . As a check we replicated the simulations for these two sets of parameters using sample sizes of 50,000, and the resulting estimates are also plotted on the graph. All of the GG estimates are relatively close, as are the chosen and fitted EW values.

Figure 2 displays additional results from our matching studies. In Panel A we show values of the K-L distance of the GG matched to the $EW(0,1,\kappa)$ distribution for a range of values of $\kappa > 0$. These distances are similar to those in Table 1, becoming even smaller as the shape parameter approaches one, where the match is perfect. In Panel B we show the logarithms of the 5th, 50th and 95th percentiles for the same EW and matched GG distributions. The agreement is clearly excellent across the range of values of the shape parameter. In Panel C we plot the matched GG parameters, together with the line of identity. As the EW shape parameter κ increases, the matched GG κ values approach the true EW values, and the matched values of the other two parameters are also increasingly close to their assumed EW values. This increasing agreement can also be seen in Figure 1. At the far left of the graph in Panel C, where the EW parameter is approximately $\kappa = 0.28$, the matched GG value of κ is approximately zero, corresponding to the lognormal distribution. We matched the EW to the standard ($\beta = 0$, $\sigma = 1$, $\kappa = 0$) lognormal and obtained matched EW parameter values of $\beta = 3.8723$, $\sigma = 0.72845$, $\kappa = 0.27795$. The K-L distance of this EW distribution from the standard lognormal was approximately $8.4E-5$. Thus the log transformation of this EW distribution provides an excellent approximation to the standard normal. For a similar comparison between the lognormal distribution based on the normal with mean zero and variance $\pi/3$ and the standard log logistic distribution, the K-L distance was 0.01436, which is orders of magnitude larger than the values obtained from our EW-GG comparisons. The log quantiles of this EW distribution agree almost perfectly with those of a standard normal (QQ plot not shown). Panel C of Figure 2 also shows that as the EW parameter κ approaches zero, the matched GG parameter β becomes increasingly negative. In Panel D we plot the survival functions for three small EW parameter values $\kappa = 0.5, 0.4, 0.3$; the matched GG distributions are indistinguishable. For the matched GG, the results were $\beta = -1.4174, -2.4666, -4.5707$. For comparison we included the same matched GG survival curves with the three β parameters set to zero. These curves are more typical of the behavior of the GG for small values of κ (0.3589, 0.2051, 0.0387). The effect of the negative values of β is fairly dramatic.

3.2 Comparisons based on fitting

Finally we turn to an application, fitting the two distributions to the same set of data and comparing the results. Table 2 contains estimated and matched parameters together with hazard shapes and the log likelihood for distributions fitted to the MACS data from Period 2, as well as the corresponding K-L distances. The first row of the table has the fitted EW distribution, the second is the fitted GG and the third contains the parameters of the GG distribution matched to the fitted EW. The parameter estimates for the fitted and matched GG are similar to each other and to the fitted EW. Interestingly the values of the log likelihood for the fitted EW and GG distributions are very close. The KL distance of the matched GG distribution is smaller than that of the fitted, although the shape of the hazard is

different, presumably reflecting that fact that the values of (σ, κ) are close to the boundary ($\sigma = \kappa$) between the two different hazard shapes (Figure 1 of [1]). The final three rows of the table give the parameter estimates for fitted EE ($\sigma = \kappa$), fitted gamma and GG matched to the fitted EE distributions. These results are similar to those for the EW and GG. It is interesting but not particularly surprising that the GG distribution matched to the EE is not a gamma distribution, and in fact has an arc-shaped hazard ($\kappa < \sigma < 1$). This is consistent with the differences seen in Figure 1, where the parameter estimates from Table 2 are also plotted. Of course, for distributions on the boundary curves of the four hazard regions, the shape of the hazard is difficult to determine very precisely.

In Panel A of Figure 3, we show the survival and hazard functions for the fitted EW and GG distributions. The survival curves are indistinguishable, although a slight difference can be seen between the two hazard functions. In Panel B we compare fitted EE and standard gamma distributions. In this case the difference between the hazard functions is more apparent. For comparison, in Panel C we show the fitted EW and matched GG, and in Panel D the fitted EE and matched GG distribution. The actual shape of the matched GG hazard is not apparent for the length of time shown in the graph. Finally in Panel E we show relative (to those for the fitted EW) quantile differences between the matched GG and fitted EW distributions, and in Panel F we plot the curves for the K-L distance for fitted GG and GG matched to fitted EW, showing the regions of greatest difference between the two densities.

We also fitted the EGG distribution to the Period 2 data. This did not result in an improved fit; the value of the log likelihood was -801.7276, nearly identical to the values in Table 2. A profile likelihood plot involving the power parameter suggested that this fit could not readily be improved. This result is not surprising given that the fitted GG distribution basically provided an adequate description of the data [1].

An additional comparison is provided by the example considered by Mudholkar and Hutson in their original paper [4]. This is a small data set consisting of 39 observations, reporting flood rates for the Floyd River. The example is interesting because the GG estimate of the shape parameter is negative: $\hat{\beta}(SD) = 8.0246 (0.254)$, $\hat{\sigma} = 1.0291 (0.121)$, $\hat{\kappa} = -0.3097 (0.372)$. The EW parameter estimates were rather different: $\hat{\beta}(SD) = 20.1937 (15.663)$, $\hat{\sigma} = 0.4875 (0.381)$, $\hat{\kappa} = 0.1133 (0.150)$. Using their parameterization (see Section 2.2), the estimates match those of Mudholkar and Hutson: $\hat{\alpha}(SD) = 0.2323 (0.131)$, $\hat{\sigma} = 4.2433 (26.30)$, $\hat{\theta} = 77.9434 (206.6)$. The log likelihood for the GG model was -57.081, and the EW log likelihood was -57.079. Both hazards are arc-shaped. In addition, the GG fit is consistent with a lognormal distribution, with log likelihood of -57.427. Plots of the GG and EW survival functions were indistinguishable, and very close to the survival function for the lognormal.

4. DISCUSSION

We have compared the EW and GG distributions, which have the same hazard shapes for the same values of the three parameters. Using a set of representative parameter values, we have shown that for each EW distribution, there is a corresponding GG distribution that is

nearly identical. In addition, fitting these two parametric distributions to the same data produces virtually the identical fit of the survival and density functions, although the parameter estimates may differ, especially when the EW shape parameter is less than one. In the applications, even the values of the log likelihood were nearly identical.

The agreement between the two distributions in our various comparisons, both graphically and in terms of the K-L distance, is striking. We regard these results as more than a curiosity, and in fact as evidence of a close connection between the two families of distributions. To our knowledge these families and their extensions are the only two having the four standard hazard shapes, which we regard as an important property that strongly recommends these distributions for use in applications. Further investigation of the two families, which otherwise seem completely unrelated, may help provide deeper understanding of parametric families of distributions sharing this important property. Our results suggest that this may in some way be a structurally important characteristic for a family of distributions.

An advantage of the EW family is that it is easier to work with than the GG. The GG is readily available in standard statistical software packages; however the EW can be fitted in standard packages using the methods we have described [1]. A limitation of the EW distribution is that it does not actually include the gamma or lognormal distributions, although it does provide extremely good approximations. For a comprehensive review of the EW distribution and a number of generalizations, see [7].

A final conclusion from our investigations is that the four-parameter EGG family appears to be an attractive option deserving further application. The GG distribution is also a member of the generalized F (GF) distribution family. With the exception of the GG, however, the hazard behavior of the GF distribution is rather limited [8]. Thus the EGG, whose hazard behavior is the same as the GG [6], may offer a more useful extension, particularly for checking the fit of the GG.

Acknowledgments

Contract/grant sponsor: National Institute of Allergy and Infectious Diseases; contract/grant numbers: U01-AI-35042, U01-AI-35043, U01-AI-35039, U01-AI-35040, U01-AI-35041, U01-AI-35004, U01-AI-31834, U01-AI-34994, U01-AI-34989, U01-AI-34993, U01-AI-42590.

Contract/grant sponsor: National Institute of Child Health and Human Development: U01-CH-32632.

REFERENCES

1. Cox C, Chu H, Schneider MF, Muñoz A. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*. 2007; 26:4352–4374. [PubMed: 17342754]
2. Mudholkar G, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*. 1993; 42:299–302.
3. Marshall, AW.; Olkin, I. *Life Distributions*. Springer-Verlag; New York: 2007.
4. Mudholkar G, Hutson A. The exponentiated Weibull family: some properties and a flood data application. *Communications in Statistics – Theory and Methods*. 1996; 25:3059–3083.
5. Pawitan, Y. In *All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, Inc.; New York: 2001.

6. Cordeiro GM, Ortega EMM, Silva GO. The exponentiated generalized gamma distribution with application to lifetime data. *Journal of Statistical Computation and Simulation*. 2009; 81:827–842.
7. Nadarajah S, Cordeiro GM, Ortega EMM. The exponentiated Weibull distribution: a survey. *Stat Papers*. DOI 10.1007/s00362-012-0466-x.
8. Cox C. The generalized F distribution: an umbrella for parametric survival analysis. *Statistics in Medicine*. 2008; 27:4301–4312. [PubMed: 18407568]

Figure 1.

A plot of selected EW parameter values (σ , κ) for each of four hazard types, together with fitted (based on simulations) and matched GG parameter values ($\hat{\sigma}, \hat{\kappa}$) ($\hat{\beta}$ not shown). Gray squares: selected EW parameters, in all cases $\beta = 0$; Triangle, diamond, circle: GG parameters estimated from simulated EW samples of sizes 200, 800 and 3200 respectively; Square: GG matched to simulated EW; $\times$: estimated GG parameters from simulated EW samples of size 50,000 (lower two points); Inverted triangles: EW parameter estimates from EW fitted to simulated EW samples of size 200. Also shown are small solid circles representing the parameters in Table 2 for EW, EE, GG and gamma distributions fitted and matched to Period 2 data.

Figure 2.

Plots of matched GG parameters for a range of values of the EW shape parameter $\kappa > 0$.

Panel A: Values of the KW distance of the GG matched to the indicated $EW(0,1,\kappa)$ distribution. Panel B: Logarithms of the 5th, 50th and 95th percentiles for the same EW (+) and matched GG (x) distributions. Panel C: matched GG parameters β (dash-dot), σ (dot), and κ (dash) together with the line of identity (solid). Panel D: Survival functions for selected EW distributions (short dash $\kappa = 0.5$, medium dash $\kappa = 0.4$, long dash $\kappa = 0.3$). Matched GG curves are indistinguishable. Shaded grey lines are the same matched GG survival functions with $\beta = 0$.

Figure 3.

EW, GG, EE and gamma distributions fitted to the data from Period 2, as well as GG parameters matched to the EW and EE distributions. Panel A: Survival and hazard functions for the fitted EW (solid) and GG (dash) distributions. Vertical reference lines show 75th and 99th percentiles of the fitted EW (solid) and GG (dash) distributions. Dotted horizontal reference lines are at 0.25 and 0.01. Panel B: Fitted EE (solid) and standard gamma (dash) distributions. Vertical reference lines show 75th and 99th percentiles of the fitted EE (solid) and gamma (dash) distributions. Solid horizontal reference lines are asymptotes of the two

hazard functions. Dotted horizontal reference lines are at 0.25 and 0.01. Panel C: Fitted EW (solid) and matched GG (dash – not seen) distributions. Vertical reference lines show 75th and 99th percentiles of the fitted EE (solid) and GG (dash – not seen) distributions. Dotted horizontal reference lines are at 0.25 and 0.01. Panel D: Fitted EE and matched GG distributions. Solid horizontal reference line is the asymptote of the EE hazard. Vertical reference lines show 75th and 99th percentiles of the fitted EE (solid) and GG (dash) distributions. Dotted horizontal reference lines are at 0.25 and 0.01. Panel E: Relative (to those for the fitted EW) quantile differences between the matched GG and fitted EW distributions. Vertical reference lines are at 0.05, 0.25, 0.50, 0.75 and 0.95. Panel F: K-L distance curves for fitted GG (dash-dot) and GG (dash) matched to fitted EW distribution.

Table 1

Comparison of GG with EW by simulation and matching. Simulation parameter estimates are means and standard deviations of fitted GG values from 1000 samples of three different sizes from the true EW distribution. KLD(GG) is the K-L distance from the fitted GG to the true EW; KLD(EW) is the distance from the fitted EW distribution.

EW	Parameters			<i>n</i>	200	800	3200	Quantile
β	σ	κ	$h(t)$	GG	Simulation	Simulation	Simulation	Match
0	.25	.75	I β	(SD)	−0.0877 (0.030)	−0.0875 (0.015)	−0.0868 (.0073)	−0.0871
				σ	0.2582 (0.016)	0.2598 (.0078)	0.2600 (.0039)	0.2601
				κ	0.6990 (0.179)	0.6981 (0.087)	0.7024 (0.043)	0.7003
				KLD (GG)	7.187E-5	1.559E-5	1.319E-5	1.287E-5
				KLD (EW)	4.001E-3	2.037E-4	1.184E-5	
0	.25	1.5	I		0.0502 (0.035)	0.0483 (0.017)	0.0483 (0.008)	0.0460
					0.2386 (0.022)	0.2413 (0.010)	0.2418 (0.0052)	0.2433
					1.6002 (0.261)	1.5678 (0.119)	1.5646 (0.059)	1.5444
					1.722E-4	2.230E-5	1.613E-5	2.529E-5
					1.091E-4	5.444E-6	1.250E-6	
0	.75	.50	A		−1.063 (0.095)	−1.062 (0.046)	v1.059 (0.023)	−1.0631
					0.8437 (0.046)	0.8489 (0.023)	0.8503 (0.011)	0.8511
					0.3588 (0.175)	0.3621 (0.087)	0.3679 (0.043)	0.3589
					1.251E-4	5.529E-5	4.989E-5	5.506E-5
					3.539E-2	1.789E-3	1.023E-4	
0	.75	2.0	B		0.1624 (0.108)	0.1585 (0.052)	0.1539 (0.025)	0.1539
					0.7129 (0.072)	0.7206 (0.034)	0.7252 (0.017)	0.7267
					2.1255 (0.291) 5.096E-4	2.0878 (0.138) 2.350E-4	2.0685 (0.066) 1.880E-4	2.0652 1.865E-4
					4.836E-5	1.838E-5	1.898E-5	
0	1.25	.50	A		−1.7723 (0.159)	−1.7699 (0.077)	−1.7655 (0.039)	−1.7718
					1.4062 (0.077)	1.4149 (0.038)	1.4171 (0.019)	1.4184
					0.3588 (0.175)	0.3621 (0.087)	0.3679 (0.043)	0.3589
					1.246E-4	5.513E-5	4.987E-5	5.505E-5
					3.524E-2	1.790E-3	1.171E-4	
0	1.25	2.0	B		0.2706 (0.180)	0.2642 (0.087)	0.2564 (0.042)	0.2565
					1.1882 (0.121)	1.2009 (0.056)	1.2086 (0.028)	1.2112
					2.1255 (0.291)	2.0878 (0.138)	2.0685 (0.066)	2.0652
					5.097E-4	2.351E-4	1.878E-4	1.862E-4
					5.686E-5	1.248E-4	2.560E-5	
0	2.0	.75	D		−0.7013 (0.241)	−0.7002 (0.116)	v0.6947 (0.059)	−0.6966
					2.0660 (0.126)	2.0782 (0.062)	2.0804 (0.031)	2.0809
					0.6990 (0.179)	0.6981 (0.087)	0.7024 (0.043)	0.7003
					6.845E-5	1.597E-5	1.322E-5	1.287E-5
					4.263E-3	2.131E-4	1.410E-5	
0	2.0	1.5	D		0.4020 (0.283) 1.9089 (0.175)	0.3865 (0.133) 1.9308 (0.082)	0.3868 (0.066) 1.9344 (0.041)	0.3678 1.9466

EW	Parameters				<i>n</i>	200	800	3200	Quantile
β	σ	κ	$h(t)$	GG	Simulation	Simulation	Simulation	Simulation	Match
					1.6002 (0.261)	1.5678 (0.119)	1.5646 (0.059)		1.5444
					1.690E-4	2.162E-5	1.605E-5		2.529E-5
					1.348E-4	7.457E-6	1.353E-6		

Table 2

Fitted EW and EE, fitted GG and gamma, and matched GG. Data are from a study of time from a diagnosis of AIDS to death in the MACS.

Parameters Distribution	β (SD)	σ (SD)	κ (SD)	Hazard Shape	Log Likelihood	K-L Distance
EW fitted	0.7704 (0.093)	0.8363 (0.039)	0.8606 (0.108)	I	-802.067	--
GG fitted	0.6493 (0.060)	0.8475 (0.048)	0.8510 (0.127)	I	-802.161	2.9787E-5 (to EW)
GG matched to EW	0.6443	0.8519	0.8367	A		3.6278E-6 (to EW)
EE ($\kappa = \sigma$) fitted	0.7864 (0.042)	0.8416 (0.026)	0.8416	I	-802.083	--
GG ($\kappa = \sigma$) fitted	0.6483 (0.041)	0.8483 (0.024)	0.8483	I	-802.161	1.7150E-4 (to EE)
GG matched to EE	0.6360	0.8600	0.8138	A		4.7669E-6 (to EE)