What does the evidence show about the shape of mortality at the oldest ages?

Dennis Feehan*

April 24, 2014

*Office of Population Research, Princeton University
dfeehan@princeton.edu, Wallace Hall, Second Floor, Princeton, NJ 08544

* * * DRAFT: PAA 2014 extended abstract, please do not cite * * *

Abstract

Many important theoretical and policy questions require an understanding of the trajectory of death rates at advanced ages. Unfortunately, studying mortality at advanced ages can be very difficult: since absolute numbers of deaths are typically very small, we must be careful to distinguish empirical regularities from stochastic noise. In this abstract, I present an analysis of the best available data on mortality above age 80 with the aim of understanding what we can conclude about the shape of the hazard function over that range. I also review and evaluate the options available for assessing the quality of a hazard model's fit to data. This problem has been discussed at length in the literature, and is of importance to researchers and policymakers in a host of fields. I expand on previous analyses by adding new data, considering more functional forms, and improving the strategy for choosing which model provides the best fit.

1 Introduction

In the developed world, life expectancy at birth continues to rise, fertility levels are low, and populations are aging at remarkable rates. (Kinsella et al., 2005; Martin and Preston, 1994). This has dramatic implications for the needs societies will face in the future. In order to fully contend with these implications, scholars and policymakers require a quantitative understanding of death rates above age 80; that is, they need a mathematical function that summarizes the age pattern of death rates at advanced ages[1] with a few parameters. This is important for several reasons.

First, the continuing debate about the limits, or lack of limits, to human longevity requires a solid understanding of the empirical evidence on death rates at advanced ages to formulate and evaluate theories that describe the process of aging (Horiuchi and Wilmoth, 1998; Thatcher et al., 1998). Many of the existing theories imply that some functional forms should reproduce observed death rates at advanced ages better than others (Gavrilov and Gavrilova, 1991; Steinsaltz and Wachter, 2006; Vaupel et al., 1979). Furthermore, a host of empirical studies have fit parametric hazard models to observational and experimental data on old-age mortality as part of an effort to evaluate and develop theories of aging.

Second, many policymakers and planners need accurate, mathematical summaries of death rates at advanced ages in order to produce forecasts and projections (Bongaarts, 2005; Tabeau et al., 2001). Here, a low-dimensional summary of death rates at advanced ages is useful because it permits projections or forecasts to focus on few parameters.

Finally, an understanding of old-age survival patterns is important for a number of other research questions of great relevance to sociology, economics, and public policy. One example is the relationship between improvements in old-age survival and changes in savings and investment behavior; understanding how savings behavior changes as a function of improvements in survival past retirement age has dramatic implications for public policy (Sheshinski, 2007). In order to build models of behavioral responses to improvements in survival at older ages, a mathematical function that captures the essential dynamics of those changes is required.

Many functional forms have been proposed to capture the shape of mortality at advanced ages (Tabeau et al., 2001; Thatcher et al., 1998). In this article, I present an analysis of the

[1] In this paper, I use the term 'advanced ages' to refer to ages over 80.

best available data on mortality above age 80, with the aim of updating and expanding our understanding of which of the functional forms commonly used to parameterize survival fit high-quality data on death rates over the age of 80 the best. Over a decade ago, Thatcher et al. (1998) studied the fit of several functional forms to a unique and carefully constructed database of deaths at advanced ages. My analysis is similar in spirit to that exercise, with a few important differences: first, I make use of the considerable volume of data that have become available since 1990. I also benefit from a recent, thorough review of the data's quality (Jdanov et al., 2008), which allows me to focus on the country-years where quality is least likely to be a problem. The quantity of high-quality data available to today is therefore considerably greater than was available a decade ago. I also expand the focus of Thatcher et al. (1998) to several new functional forms that have appeared in the literature since their study.

Finally, I make some modifications to the methods applied in Thatcher et al. (1998); in particular, I propose the use of more principled criteria for assessing goodness-of-fit across the functional forms considered. These criteria include an appropriate penalty for the increased complexity of functional forms with more parameters. This is very important, as it provides us with some guard against over-fitting idiosyncracies in the data that might result from small sample sizes at advanced ages.

The full paper will explore the theoretical implications of these initial, empirical results, and will be accompanied by an R package to allow others to perform similar investigations in the future.

2 Methods

Data

I use the Kannisto-Thatcher (K-T) database on Old Age Mortality, which is the best collection of data on mortality at advanced ages available (Kannisto, 1994; MPIDR, 2014). The K-T database contains data on deaths and exposure above age 80 for 35 countries, with some of the Scandinavian data going back to the mid-18th century. However, data quality for mortality at advanced ages is a serious concern; see Jdanov et al. (2008) for a thorough summary of the problems involved in analyzing them. In order to ensure that poor-quality data do not mislead us in selecting useful functional forms, I only study country-years of data that were found to be of acceptable quality by Jdanov et al. (2008). Specifically, I retained the cohort data from from all countries where more than half of the years were of the highest quality, and the remaining years were of the second-highest quality in the assessment the authors provided. Furthermore, I only selected countries and time periods where deaths were reported by single year of age up to 104. This leaves us with data from Denmark, France, West Germany, Italy, Japan, the Netherlands, Sweden, and Switzerland, making a total of 285 country-years of data for each sex and single year of age. Table 1 describes the years and cohorts in our dataset for each country.

	cohort start	cohort end
Denmark	1841	1896
France	1871	1892
W. Germany	1881	1896
Italy	1881	1896
Japan	1871	1896
Netherlands	1871	1896
Sweden	1821	1896
Switzerland	1850	1896

Table 1: Cohorts from the Kannisto-Thatcher Database on Old Age Mortality used in this analysis. Cohort data are from all possible years for the countries Jdanov et al. (2008) found to be acceptable in their systematic review, and for which deaths by single year of age are available (not imputed) up to age 104; that is, I choose all cohorts from countries that Jdanov et al. (2008) found to have high-quality data based on their systematic review of the countries' period data. I considered countries to have high quality data when Jdanov et al. (2008) found that more than half of the years were of highest quality, and the remaining years were of the second-highest quality. I model ages 80-104.

Hazards

The central object of study in this project is the mortality hazard function. If the survival curve for our population of interest is $S(z)$, with $S(z) \in [0, 1]$ for all ages z, then the hazard is given by

$$\mu(z) = -\frac{d \log S(z)}{dz} = -\frac{1}{S(z)} \frac{dS(z)}{dz}. \tag{1}$$

This is the negative log-derivative of the survival function or, intuitively, it is the instantaneous rate of change of survival at age z divided by the fraction of the population left at age z, and multiplied by -1. Hazard functions are nonnegative, since survival by age can never increase and so $\frac{dS(z)}{dz}$ is always less than or equal to 0, and $S(z)$ is always between 1 and 0.

In practice, we are not able to observe the actual hazard function; instead, we often compute central death rates as an approximation. These are given by

$$M_z = \frac{D_z}{N_z}, \tag{2}$$

where D_z is the number of deaths between ages z and $z+1$ in the time period being considered [2], and N_z is the number of person-years of exposure in the time period being considered. The central death rate we observe is a function of the continuous, underlying hazard of death function $\mu(z)$.

For cohort data, we can also directly observe the probability of dying before age $z+1$, conditional on surviving to exact age z. This is given by

$$\pi(z, z+k) = \frac{D_z}{S_z}, \tag{3}$$

where D_z is again the number of deaths betwen exact ages z and $z+1$, and S_z is the number of members of the cohort who survive to exact age z.

The continuous hazard $\mu(z)$ is related to the probability of death between exact ages z and $z+k$ through

$$\pi(z, z+k) = \exp\left(-\int_z^{z+k} \mu(x)dx\right). \tag{4}$$

For more details on hazard functions and the technical study of mortality in general, see Preston et al. (2001) or Keyfitz et al. (2005).

[2] It may be helpful to recall that deaths between exact ages z and $z+1$ are the same as deaths to people whose age last birthday was z.

3 Model

I assume that cohort deaths between ages z and $z+1$ are distributed binomially; that is, if N_z people from a cohort survive to exact age z, and all of the members of the cohort face the same hazard $\mu(z)$, then

$$D_z \sim \text{Binomial}(N_z, \pi(z, z+1)), \tag{5}$$

where D_z is the number of deaths between ages z and $z+1$, and $\pi(z, z+1)$ is the probability of dying between ages z and $z+1$. $\pi(z, z+1)$ is derived from the hazard function according to Equation 4 above. This is the same model used in Thatcher et al. (1998).

The likelihood for an observed sequence of deaths $\mathbf{D} = D_1, D_2, \ldots$ and survivors to each age, $\mathbf{N} = N_1, N_2, \ldots$ is then

$$Pr(\mathbf{D}|\mathbf{z}, \theta, \mathbf{N}) = \prod_z \binom{N_z}{D_z} \pi(z, \theta)^{D_z} (1 - \pi(z, \theta))^{(N_z - D_z)}.$$

Taking logs and dropping terms that don't vary with $\pi(z, \theta)$, I have

$$ll(\mathbf{D}|\mathbf{z}, \theta, \mathbf{N}) \propto \sum_z \left[D_z \log(\pi(z, \theta)) + (N_z - D_z) \log(1 - \pi(z, \theta)) \right]. \tag{6}$$

This is the likelihood I maximize, as a function of θ, in order to fit each hazard function to the data.

Some previous work has adopted strategies for fitting mortality models that are based on fitting curves to estimated central death rates using least squares or one of its weighted or nonlinear variants. Although this approach has the advantage of being comparatively simple, there are several reasons to prefer the maximum likelihood approach I have adopted here. **?** illustrates some of the problems that arise when using regression techniques to fit parametric hazard models to central death rates, and concludes that maximum likelihood estimation is the preferred technique. Pletcher (1999) also advocates for a maximum-likelihood approach, and provides a review of some of its advantages. Wang et al. (1998) illustrates some of the ways that central death rates can be misleading approximations of

name	parameters	function	sample reference
Gompertz	α, β	$\mu(z) = \alpha \exp(\beta z)$	Gompertz (1825)
Kannisto	α, β	$\mu(z) = \frac{\alpha \exp(\beta z)}{1+\exp(\beta z)}$	Thatcher et al. (1998)
Weibull	α, β	$\mu(z) = \alpha z^{\beta}$	Gavrilov and Gavrilova (1991)
Makeham	α, β, γ	$\mu(z) = \gamma + \alpha \exp(\beta z)$	Makeham (1860)
Log-Quadratic	α, β, γ	$\mu(z) = \alpha + \beta z + \gamma z^2$	Coale and Kisker (1990); Wilmoth (1995)
Logistic	α, β, γ, δ	$\mu(z) = \gamma + \frac{\alpha \exp(\beta z)}{1+\delta \exp(\beta z)}$	Beard (1971); Perks (1932)
Lynch-Brown	α, β, γ, δ	$\mu(z) = \alpha + \beta \arctan\{\gamma(z-\delta)\}$	Lynch and Brown (2001)

Table 2: The functional forms for the hazard of death at advanced ages considered in this analysis. In the functions listed, z is age, $\mu(z)$ is the force of mortality at age z. In compiling this list, I am particularly indebted to the discussion of various functional forms, and their origins in the literature, contained in Gavrilov and Gavrilova (1991) and Thatcher et al. (1998).

the underlying force of mortality at old ages.

Functional Forms

I consider several functional forms that have been employed to model mortality in adult and advanced ages; these are listed in Table 2. Each of these forms has two or more parameters; once the parameters are fixed, the shape of the hazard funtion is completely determined. The technical appendix has a more detailed explanation of each function.

Model selection

I measure the goodness of fit of the various functional forms in several ways: first, I compute the sum of squared errors in the estimated number of deaths for each country-sex-year; that is, I compute

$$\text{SSE} = \sum_z \left(\hat{D}_z - D_z\right)^2,$$

where k is the number of ages being fit. This quantity gives us a measurement of how close each model's predictions come to the observed numbers of deaths, in absolute terms.

Second, I compute Akaike's Information Criterion (AIC) for each model and use it to rank them within each country-sex-year; that is, I compute the AIC values within the country-sex-year and rank them from best to worst. The AIC is essentially a penalized log-likelihood, where the penalty is a function of the number of parameters being estimated in the model:

$$\text{AIC} = -2\mathfrak{L} + 2k, \tag{7}$$

where $\mathfrak{L}$ is the value of the maximized likelihood from Equation 6, and k is the number of parameters being estimated in the model. Although the AIC has a simple form, it can be justified as selecting the model minimizes the estimated Kullback-Leibler distance between the distribution of data implied by the model and the one seen in the data (Burnham and Anderson, 2004; Claeskens and Hjort, 2008). Absolute AIC values are not interpretable, but the ordering of models given by their AIC values is. The ranking of models by their AIC values, which I compute for each country-sex-year, is thus an indication of how well the various functional forms trade off the number of parameters estimated with the accuracy of their fit to the data. A tradeoff is necessary here because a functional form with too many parameters is at risk of overfitting – that is, if I allow too many parameters to be estimated, I risk capturing esoteric features of the dataset I happen to have observed. I would not want to evaluate theories or produce forecasts based on a model that would reproduce esoteric features of our data that would not be found in other settings.

Third, I compute the Bayesian Information Criterion (BIC) for each model, and also use it to rank the models within each country-sex-year from best worst. Although the BIC has a different theoretical justification from the AIC, it too has a simple form:

$$\text{BIC} = -2\mathfrak{L} + 2k \log k, \tag{8}$$

where $\mathfrak{L}$ is the value of the maximized likelihood from Equation 6, and k is the number of parameters being estimated in the model. In general, the BIC penalizes complex models more heavily than the AIC. There is much discussion in the statistical literature about the virtues of one of these information criteria over the other (Burnham and Anderson, 2004; Claeskens and Hjort, 2008; Kass and Raftery, 1995). There are advantages to each, and there is no conclusive reason to prefer one over the other; of course, the best situation is when they both agree on the ordering of a set of models.

4 Results

4.1 Two illustrative examples

As a sample of the results, Figure 1 shows the observed and fitted death rates from the seven functional forms for one country-sex-cohort in the dataset: the Netherlands, females, 1885. The blue circles show the observed central death rates, while the black curves show the maximum likelihood fit of each functional form. The area of each blue circle is proportional to the estimated person-years of exposure at each age. As we would expect, the observed hazard, approximated by the central death rates, appears to increase with age. There is some evidence of a plateau at the highest ages, though the greatly reduced size of the cohort by that point means that it is difficult to be certain. For this cohort, it appears as though some functional forms fit the observed death rates much better than others; in particular, the four-parameter Lynch-Brown and Logistic functions are able to bend at advanced ages to accommodate what might be a tapering in the hazard.

Measures of model fit for each hazard function shown in Figure 1 are reported in Table 3. The functional forms are ordered by the rank they attain using the AIC, where rank 1 is the model that fits the data the best and rank 10 is the model that fits the data the worst. The table also shows the sum of squared errors in the estimated number of deaths at each age (SSE), the difference between each model's AIC and the minimum AIC (ΔAIC), and the difference between each model's BIC and the minimum BIC (ΔBIC). A few things are remarkable about the results. First, comparing the AIC and BIC ranks with the SSE shows that the AIC and BIC do not uniformly prefer models whose predictions come closest to the observed deaths in absolute terms; they trade off between a model's fit and its complexity. Looking first at the SSE, we see that the four-parameter Lynch-Brown model fit the data most accurately, as we would expect. However, that the three-parameter Quadratic model achieves a fit that is very close to the Lynch-Brown model, but using only three parameters. Indeed, we see that for this cohort, both the AIC and the BIC give the Quadratic model the top ranking.

Figure 2 shows the same hazard functions fit to a second cohort, Swedish women born in 1885. This case is somewhat different: here, the central death rates increase quite smoothly with age, and all of the hazard functions except for the Weibull have more or less the same shape; in particular, the two-parameter Gompertz and Kannisto models appear to have produced very similar fits to the four-parameter Logistic and Lynch-Brown functions. Table 4 has the detailed measures of fit for this cohort. We see that, in this case, the advantages of extra parameters are less apparent. The AIC ranks the two-parameter Kannisto and Gompertz functional forms first and third; in second place is the three-parameter Quadratic function. Also, unlike the example from the Netherlands we just considered in Figure 1, for this cohort the AIC and the BIC do not completely agree on the ordering of the models; in particular, the log-quadratic model was in second place according to the AIC, but it is in third place according to BIC. This is an example of how the BIC penalizes additional parameters more severely than the AIC. Nevertheless, the rankings between the two methods are very similar to one another.

As a rule of thumb, Burnham and Anderson (2004) suggests that models with $\Delta\text{AIC} \leq 2$ have some support from the data and, similarly, Kass and Raftery (1995) suggests that the evidence against models with $\Delta\text{BIC} \leq 2$ is weak, meaning the models themselves receive support from the data. In the example from Belgium shown in Table 3, these rules of thumb suggest that, while the log-quadratic model is the top performer, the data also provide some support for the Lynch-Brown and, possibly, the Logistic models.

4.2 General results

Fits like the ones in Figures 1 and 2 were produced for each cohort in the dataset. In order to draw conclusions about how well the functional forms fit the data in general, we will now turn to aggregate measures of their performance across all of the country-sex-cohorts in our data.

The final version of the paper, still underway, will have a detailed analysis of the overall results shown in Figures 3, 4, and others.

5 Conclusion

Many studies of mortality at advanced ages, and related substantive topics, require that the analyst choose a functional form to summarize the hazard of death by age. In these studies, the analyst should employ principled criteria for selecting which functional form to use. In particular, the criteria used to select a model should explicitly address the tension between fitting the observed data as well as possible, on the one hand, and minimizing model complexity and, therefore, the risk of over-fitting, on the other.

In this preliminary analysis, I selected seven functional forms used in a wide spectrum of studies, from forecasting to theoretical modeling, and demonstrated the use of powerful, principled tools – Akaike's Information Criterion and the Bayesian Information Criterion – to select the most attractive model. Of the seven functional forms I considered, the log-quadratic model had the strongest support from a cross-national, high-quality dataset of deaths at advanced ages. Several other functional forms also performed well, but the one most commonly recommended for use in the literature, the logistic, generally produced only moderately good fits.

In many cases, there will be considerations beyond statistical ones which will limit the set of candidate models that the analyst will consider. For example, a study that seeks to evaluate the level of empirical support for theories of aging may only wish to consider functional forms suggested by one of the theories being investigated. The procedure outlined above, or a similar one, should still be employed on the reduced set of candidate models to decide which ones enjoy the strongest empirical support.

The full version of the paper will explore the results presented above in more depth, with a special focus on the implications the accuracy of the various hazard functions has for our theoretical understanding of old-age mortality.

References

Beard, R. E. (1971). Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. *Biological aspects of demography*, page 5768.

Bongaarts, J. (2005). Long-range trends in adult mortality: Models and projection methods. *Demography*, 42(1):2349.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research*, 33(2):261304.

Claeskens, G. and Hjort, N. L. (2008). *Model selection and model averaging*, volume 330. Cambridge University Press Cambridge.

Coale, A. J. and Kisker, E. E. (1990). Defects in data on old-age mortality in the United States: new procedures for calculating mortality schedules and life tables at the highest ages.

Gavrilov, L. A. and Gavrilova, N. S. (1991). *The biology of life span: a quantitative approach*. Harwood Academic Publishers, Chur [Switzerland]; New York.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115:513–583.

Horiuchi, S. and Wilmoth, J. R. (1998). Deceleration in the age pattern of mortality at olderages. *Demography*, 35(4):391412.

Jdanov, D. A., Jasilionis, D., Soroko, E. L., Rau, R., and Vaupel, J. W. (2008). Beyond the Kannisto-Thatcher database on old-age mortality: An assessment of data quality at advanced ages. *MPIDR Working Papers*.

Kannisto, V. (1994). Development of oldest-old mortality 1950-1990: evidence from 28 developed countries.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Keyfitz, N., Caswell, H., Caswell, H., and Keyfitz, N. (2005). *Applied mathematical demography*, volume 47. Springer.

Kinsella, K., Phillips, D., and Bureau, P. R. (2005). *Global aging: The challenge of success*. Population Reference Bureau.

Lynch, S. M. and Brown, J. S. (2001). Reconsidering mortality compression and deceleration: An alternative model of mortality rates. *Demography*, 38(1):7995.

Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *The Assurance Magazine, and Journal of the Institute of Actuaries*, page 301310.

Martin, L. and Preston, S. (1994). *Demography of aging*. National Academies Press.

MPIDR (2014). Kannisto-Thatcher database on old-age mortality. http://www.demogr.mpg.de/databases/ktdb/introduction.htm.

Perks, W. (1932). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, page 1257.

Pletcher (1999). Model fitting and hypothesis testing for age-specific mortality data. *Journal of Evolutionary Biology*, 12(3):430–439.

Preston, S., Heuveline, P., and Guillot, M. (2001). Demography: measuring and modeling population processes. *Population and Development Review*, 27:365.

Sheshinski, E. (2007). *The Economic Theory of Annuities*. Princeton University Press.

Steinsaltz, D. and Wachter, K. (2006). Understanding mortality rate deceleration and heterogeneity. *Mathematical Population Studies*, 13(1):1937.

Tabeau, E., Van den Berg Jeths, A., and Heathcote, C. (2001). *Forecasting mortality in developed countries*. Springer.

Thatcher, A., Kannisto, V., and Vaupel, J. (1998). *The force of mortality at ages 80 to 120*, volume 22. Odense University Press Odense.

Vaupel, J., Manton, K., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439454.

Wang, J.-L., Muller, H.-G., and Capra, W. B. (1998). Analysis of oldest-old mortality: Lifetables revisited. *The Annals of Statistics*, 26(1):126–163.

Wilmoth, J. R. (1995). Are mortality rates falling at extremely high ages? An investigation based on a model proposed by Coale and Kisker. *Population studies*, 49(2):281295.

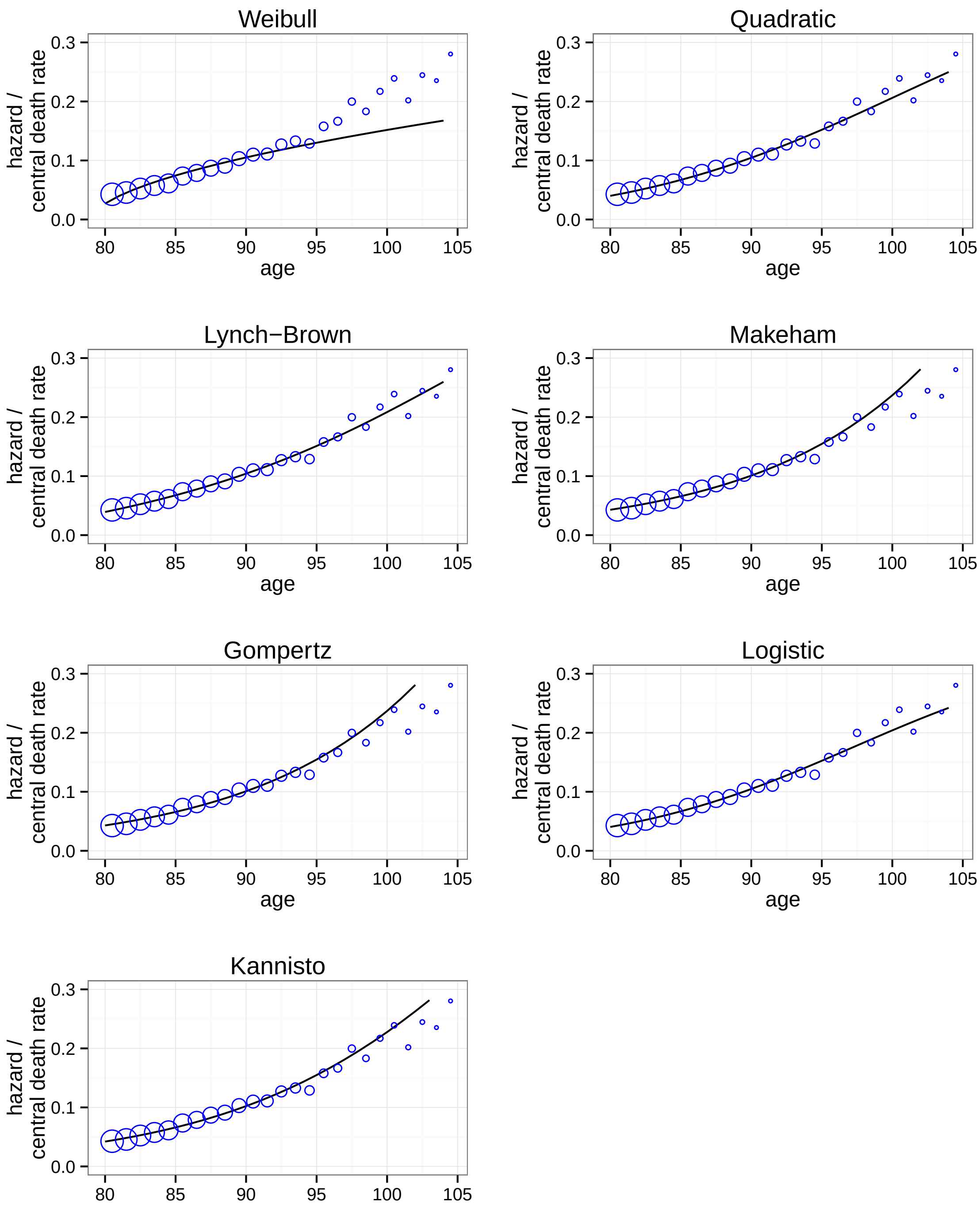

Figure 1: The functional forms fit to the cohort data for females in the Netherlands, 1885. The blue circles show the observed central death rates, while the black curves show the maximum likelihood fit of each functional form. The area of each blue circle is proportional to the estimated person-years of exposure at each age. There is some suggestion of a tapering of the central death rates at advanced ages, but small numbers of surviving chort members at those ages makes this difficult to determine graphically.For this cohort, some functional forms fit the observed death rates much better than others; in particular, the Quadratic and Lynch-Brown models appear to provide the most accurate fits. Measures of model fit for each hazard function shown here are reported in Table 3.

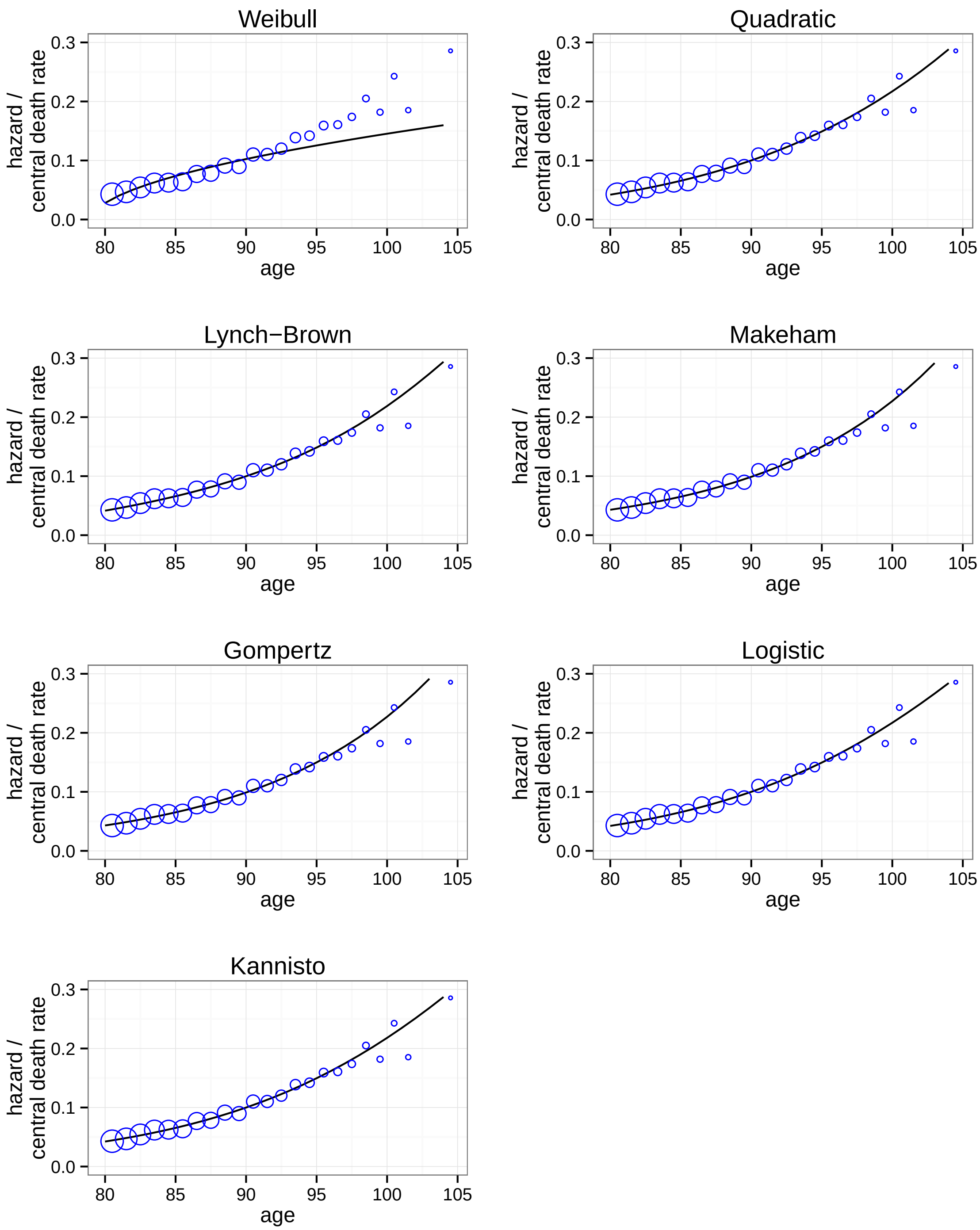

Figure 2: The functional forms fit to the cohort data for females in Sweden, 1885. The blue circles show the observed central death rates, while the black curves show the maximum likelihood fit of each functional form. The area of each blue circle is proportional to the estimated person-years of exposure at each age. The AIC rankings and sum of squared error (SSE) values for this country-sex-year are shown in Table 4. For this country-sex-year, some functional forms fit the observed death rates much better than others. In particular, the AIC suggests that the Kannisto and Quadratic models fit the data the best.

	log-likelihood	SSE	AIC	AIC rank	Δ AIC	BIC	Δ BIC	BIC rank	CV RMSE(Dx)	CV ran
Quadratic	-68022.24	17590.54	136050.49	1	0.00	136056.15	0.00	1	2616.73	2
Lynch-Brown	-68021.70	17590.36	136051.40	2	0.91	136058.27	2.12	2	2611.00	1
Logistic	-68023.05	17590.73	136054.10	3	3.61	136060.97	4.83	3	2661.51	3
Kannisto	-68030.07	17591.77	136064.14	4	13.65	136068.57	12.43	4	3222.79	4
Gompertz	-68037.56	17593.24	136079.12	5	28.63	136083.56	27.41	5	3758.28	5
Makeham	-68037.56	17593.24	136081.12	6	30.63	136086.78	30.63	6	3758.36	6
Weibull	-68137.87	17608.53	136279.73	7	229.24	136284.17	228.02	7	10858.59	7

Table 3: Measurements of model fit for the cohort of Dutch females born in 1885.

	log-likelihood	SSE	AIC	AIC rank	Δ AIC	BIC	Δ BIC	BIC rank	CV RMSE(Dx)	CV ran
Kannisto	-59229.23	15286.05	118462.45	1	0.00	118466.89	0.00	1	3665.80	1
Quadratic	-59228.98	15286.02	118463.96	2	1.51	118469.62	2.73	3	3702.08	3
Gompertz	-59230.46	15286.23	118464.92	3	2.47	118469.36	2.47	2	3714.04	4
Lynch-Brown	-59228.60	15285.98	118465.19	4	2.74	118472.07	5.18	4	3666.13	2
Logistic	-59229.23	15286.06	118466.46	5	4.01	118473.34	6.45	6	3714.53	6
Makeham	-59230.46	15286.23	118466.92	6	4.47	118472.58	5.69	5	3714.12	5
Weibull	-59358.43	15308.92	118720.86	7	258.41	118725.30	258.41	7	10709.70	7

Table 4: Measurements of model fit for the cohort of Swedish females born in 1885.

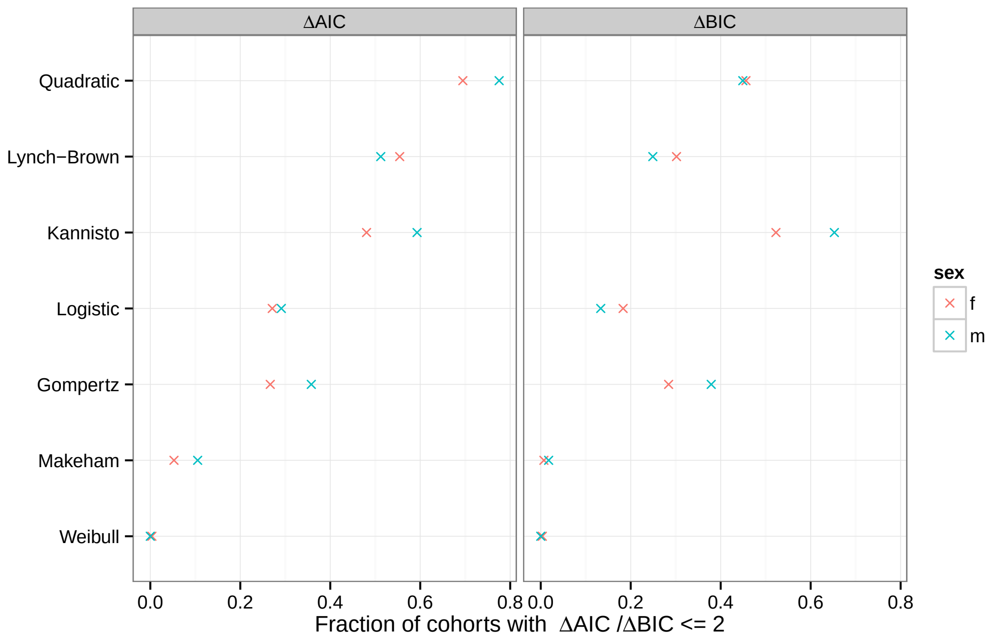

Figure 3: Preliminary results summarized by functional form and fit measure. As explained in the text, if ΔAIC/BIC < 2 then, as a rough rule of thumb, the model can be considered to fit the data reasonably well. This plot shows the fraction of the cohorts studied for which each functional form met that criterion; on the left-hand panel, we see the results for AIC, and the right-hand side shows the results for BIC. Results for males are in red, and females are in blue.

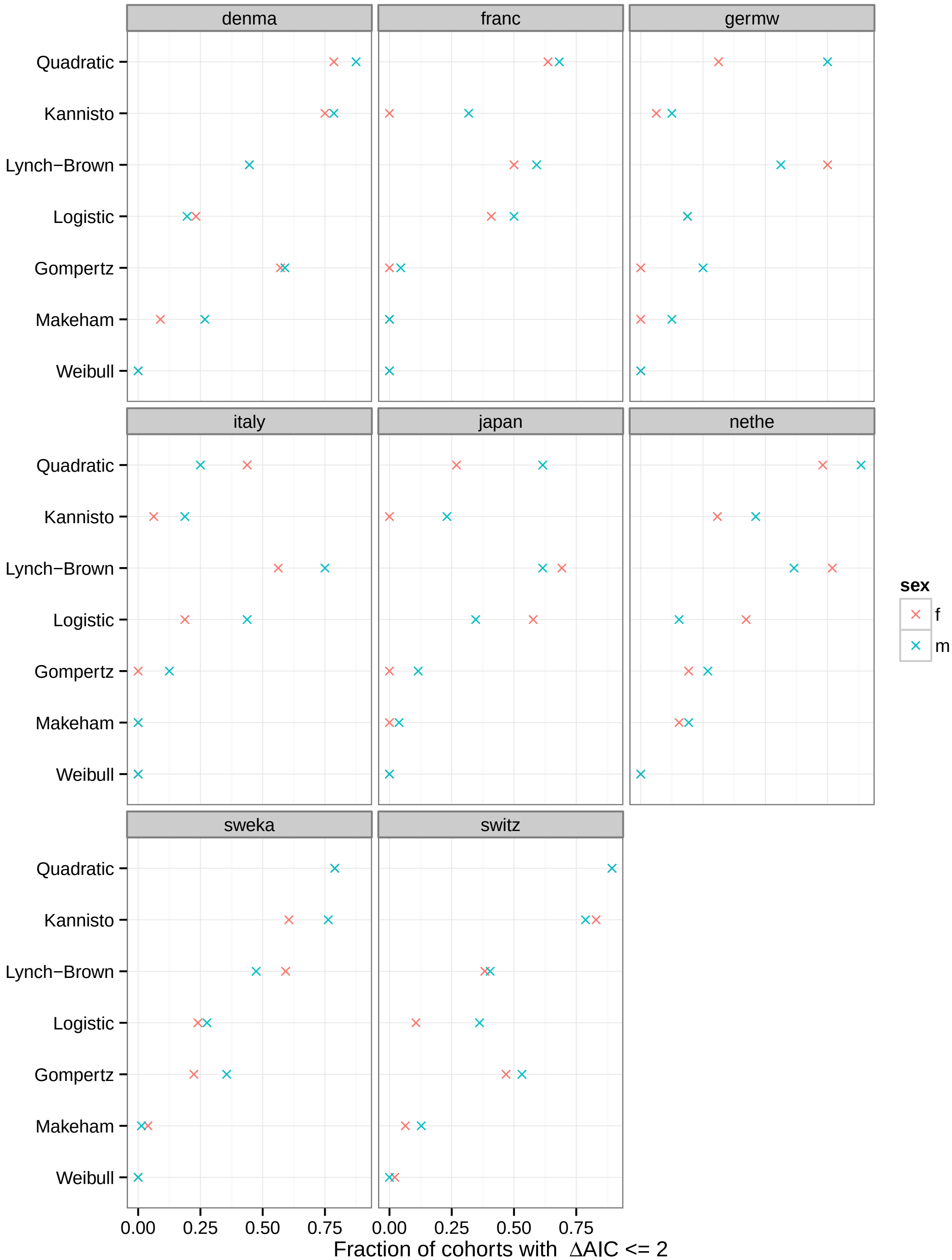

Figure 4: Preliminary results for the fit measure ΔAIC, summarized by functional form and country. This is the same data shown in the left-hand panel of Figure 3, now broken down by country. Results for males are in red, and females are in blue. We can see that there is considerable variation across countries in which hazards appear to fit best.