Using Parametric Methods to Extrapolate Survival beyond Trial Period with Application to Long Term Cost Effectiveness analysis

Guiqing Yao[a], Nick Freemantle[a], Pelham Barton[a], John Cleland[b], Stirling Bryan[a]

[a] University of Birmingham, UK.

[b] Department of Cardiology, University of Hull, Castle Hill Hospital, Kingston upon Hull, UK.

Address for Correspondence

Mrs Guiqing Yao

Health Service Management Center

The University of Birmingham

40 Edgbaston Park Road

Birmingham UK

B15 2RT

Tel: +44 (0)121 414 3197

Fax: +44 (0)121 414

Email: G.L.YAO@bham.ac.uk

Abstract

Aims: In this paper, take the common situation where the cost-effectiveness analyst wishes to extrapolate beyond the trial but does not have access to the individual patient data. We describe the most appropriate methods for extrapolation. We aim to provide an overall review of the methodology and a practical guide.

Methods: First, we review three most commonly used models in survival analysis, stressing the characteristics of their hazard functions. Second, we describe the methods for choosing a suitable parametric model and assessing its goodness-of-fit. Third, we illustrate this method using data from a recently published trial. Fourth, we compare the result of the parametric fitting by using individual data with the one derived from published Kaplan Meier estimates.

Data: the CARE HF trial

Results: The derivation of an appropriate functional form for the parametric survival function based on published Kaplan Meier curve using our method was a robust estimation compared with the one derived from individual data. However the variance based on Kaplan Meier survival function was underestimated if no individual variation were considered. The variance should be inflated by 30% to reflect the true variance in the case of the CARE HF data.

Conclusions: An estimate of uncertainty is crucial for economic analysis, where it is quite inappropriate to take the observed survival function and assume that it has been measured without error. Variance estimation should account for both variation along the Kaplan Meier curves and variation in the reported Kaplan Meier estimates.

1 Introduction

It is widely accepted that randomised controlled trials provide the best evidence of the effect of health care interventions. However, the economic consequences of health care interventions are rarely described fully by randomised trials, in part because ethically it is not possible to continue a trial beyond the point at which effectiveness has been established, even though issues of cost-effectiveness may still be uncertain. Thus, the requirement for longer-term final outcomes for economic evaluations (e.g. life years gained and quality-adjusted life years gained) often makes it necessary to extrapolate cost and effectiveness beyond the period observed in a trial, using modelling methods. Survival time is, therefore, often a key factor in cost-effectiveness analyses and results tend to be sensitive to the estimate of this parameter.

Most survival analysis in the current medical literature uses non-parametric methods. The Kaplan-Meier approach gives an estimate of survival time for the observed follow-up only; in order to consider survival beyond the observed data, parametric models can be employed as extrapolations of survival possible functions. Whilst exponential and Weibull distributions are frequently used for this purpose, consideration of the appropriate distribution, given the observed survival data, has received little attention in the health economics literature. Little work has been done on how to choose a family of distributions based on observed data and how to evaluate the appropriateness of the chosen distributions.

In this paper, take the common situation where the cost-effectiveness analyst wishes to extrapolate beyond the trial but does not have accessed to the individual patient data. We describe the most appropriate methods for extrapolation, using a range of parametric methods, including methods for the robust derivation of survival functions from published Kaplan Meier curves. We provide an overall review of the methodology and a practical guide. The paper organized as follows: first, we reviewed three most commonly used models in survival analysis, stressing the characteristics of their hazard functions; second, we describe the methods for choosing a suitable parametric model and assessing its goodness-of-fit; and third, we illustrate this method using data from a recently published trial (CARE HF).

2 Parametric survival functions

Many probability distributions correspond to some specific process in the medical field. Each of those distributions has their unique characteristics in their hazard functions. One way to select parametric distributions for a particular survival time might be based on the characteristics of their hazard functions.

For example, an exponentially distributed survival time corresponds to the assumption of a constant hazard, while the Weibull model implies a monotone hazard. For log-logistic and lognormal distributions, the hazard functions increase first; reach a peak and then decrease. A Gompertz hazard implies short term increasing or decreasing, then constant thereafter as implied in an exponential distribution. Lognormal and Gamma distributions provide a wider range of survival functions.

In this section, we review three parametric distribution functions used in survival analysis literature and describe their hazard functions.

2.1 Parametric survival functions

2.3.1 Exponential distribution

The exponential distribution is commonly used in health economic evaluation. However when survival time is assumed to follow an exponential distribution, the implication is that the hazard function is constant.

For an exponential distribution, the hazard function is

$$h(t) = \lambda \qquad (2.1.1a)$$

The corresponding survival function is

$$S(t) = \exp(-\lambda t) \qquad (2.1.1b)$$

The limitation of the exponential distribution is the constant hazard property. It means that the probability that an event occurs for the next time period given survival at the current time does not depend on the patient's history. As time progresses for a particular individual, the (conditional) probability of death in successive time intervals remains unchanged. This, however, is not plausible in most clinical settings.

2.3.2 Weibull function

The Weibull distribution is another frequently used distribution in health economic modelling. There are two parameters in a Weibull distribution, the scale parameter λ and the shape parameter γ. The hazard function for Weibull survival time could be increasing or decreasing with time depending on the shape parameter γ. If the shape parameter is

greater than 1, the hazard rate increases with time. If the shape parameter is less than 1, the hazard decreases with time. If the shape parameter is equal to 1, then the Weibull reduces to the exponential distribution.

2.3.3 Hazard function for Weibull model

2.3.4 $h(t) = \lambda\gamma t^{\gamma-1}$ *(2.1.2a)*

Survival function for the model

$$S(t) = \exp(-\lambda\ t^{\gamma}) \qquad (2.1.2b)$$

From 2.1.2a, we can see that if $\gamma = 1$, the hazard function $h(t)$ simplifies to the constant value λ, which is the hazard function for the exponential distribution.

One limitation of the Weibull distribution is that the hazard function is a monotonic function of time. However, in medicine, it is not uncommon to see a hazard rate changes direction over time. Therefore alternative distributions are needed for those circumstances.

2.3.5 The log-logistic distribution

A standard log-logistic model has two parameters λ and β. If survival time follows a log-logistic distribution then its logarithm has a logistic distribution. The hazard function of a log-logistic distribution has a single peak. This could represent a typical clinical process of a disease treatment. For example, following a renal transplantation, a patient faces an increasing hazard of death over the first few months after the transplant, the hazard then decreases with time as the patient adapted to the new graft.

Hazard function for log-logistic model

$$h(t) = \frac{\lambda\beta\ t^{\beta-1}}{1+\lambda t^{\beta}} \qquad (2.1.3a)$$

The hazard function decreases monotonically if β <=1, but if β >1, the hazard has a single peak.

The survival functions of log logistic distribution:

$$S(t) = [1+\lambda t^{\beta}]^{-1} \qquad (2.1.3b)$$

The advantage of log logistic hazard function is that it captures both inverted U-shaped and monotonically declining rates.

2.2 Choice of parametric models

Prior to fitting a parametric model, a preliminary study of suitability of an assumption should be carried out. Most studies in clinical trials usually report time to event survival analysis by using Kaplan-Meier estimates. Once the Kaplan-Meier estimates of survival function have been obtained, the raw hazard rate for a given period can be calculated. A hazard rate at a given time period can be calculated by the following method:

$$h(t) = \frac{S_t - S_{t-\Delta t}}{S_{t-\Delta t} * \Delta t} \quad (2.2)$$

A plot of the estimated hazard rate against time may then suggest a suitable parametric form for the hazard function.

Figure 1 shows four examples where hazard function based on equation (2.2) was plotted against time. Figure 1a shows slightly decreasing hazards but almost constant, so a Weibull or exponential distribution should be considered. Figure 1b has a decreasing trend with time; Weibull or log logistic could fit the data well. Figure1c, a Weibull with

increased hazard function should be attempted and figure 1d indicated initial decreasing hazards and reach lower points then increasing with time - a Gamma distribution could be considered.

2.3 Parameter estimates of a chosen survival function

Once a family distribution has been selected, the next step is to estimate the parameters of the distribution.

2.3.1 Exponential distribution

Exponential survival function with one parameter is shown below:

$$S(t) = \exp(-\lambda t) \qquad (2.4.1a)$$

Take logarithm on both side leads to:

$$\log(S(t)) = -\lambda t \qquad (2.4.1b)$$

If we have an estimate of the survival function at a range of times, we read a set of data points (t_i, S_{t_i}) where i= 1, 2, …, k.

Now fitting the model

$$\log(S(t_i)) = -\lambda t_i + \varepsilon_i, \quad (2.4.1c)$$

where ε_i is an error term. We can use least squares methods to estimate *λ, and hence the* survival function at any given time.

2.3.2 Weibull distribution

Suppose that a Weibull distribution for the survival times is attempted. For a Weibull distribution with scale parameter λ and shape γ, the survival function is given by

$$S(t) = \exp(-\lambda t^{\gamma}) \qquad (2.4.2a)$$

Let us linearise the above equation. Taking the logarithm on both sides,

$$\log(S(t)) = -\lambda t^{\gamma} \qquad (2.4.2b)$$

Take another logarithm on both sides again:

$$\log(-\log(S(t)) = \log(\lambda) + \gamma \log(t) \qquad (2.4.2c)$$

Now we fit a model based on the relationship above:

$$y_i = k + \gamma x_i + \varepsilon_i \qquad (2.4.2d)$$

Where ε_i is an error term and

$$y_i = \log(-\log(S(t_i)) \qquad (2.4.2e)$$

$$x_i = \log(t_i) \qquad (2.4.2f)$$

$$k = \log(\lambda) \qquad (2.4.2g)$$

We can use equations (2.4.2e) and (2.4.2f) to transform data points (t_i, S_{t_i}) into pairs (x_i, y_i). we then uses least squares methods in model (2.4.2d) to estimate the two parameters k and γ with their confidence intervals. Putting $\lambda = \exp(k)$ gives us the necessary parameters for the Weibull distribution.

2.3.3 Log-Logistic model

For the log-logistic survivor function for this distribution is

$$S(t) = [1 + \lambda t^{\beta}]^{-1} \qquad (2.4.3a)$$

We can rearrange this to give us

$$\frac{1 - S(t)}{S(t)} = \lambda t^{\beta} \qquad (2.4.3b)$$

And hence

$$\log(\frac{1-S(t)}{S(t)}) = \log(\lambda) + \beta\log(t) \qquad (2.4.3c)$$

We now have a linear model to which to apply methods similar to those described above for the Weibull distribution.

2.5 Model checking

Whatever parametric model is fitted, it is always necessary to examine the goodness of fit. Statistical tests, graphical methods or examining the residuals are commonly used approaches for goodness-of-fit testing in survival analysis. When we only have read data from Kaplan Meier estimates without accessing individual data, the most appropriate choice would be based on graphical methods.

An informative way of assessing whether a particular distribution represents a survival time is to compare the survivor function for the data with that of a chosen model. Plot of fitted survival function with Kaplan Meier survival curves would give a visual evidence of modelling fitting. In addition, hazard plots from a raw calculation compared with fitted hazard functions with corresponding confidence interval would provide more information about the fit.

3 Data: the CArdiac REsychronisation in Heart Failure (CARE-HF) trial

The design and results of the CARE-HF study have been reported previously. A total of 813 patients were randomly assigned to receive medical therapy (MT) alone (404) or with a cardiac resychronisation device therapy (CRT) plus MT (409). The mean duration of follow-up was 29.4 months (range 18.0–44.7). The primary endpoint was the time to death from any cause or unplanned hospitalization for a major cardiovascular event. The

principal secondary endpoint was death from any cause. The cost effectiveness analysis was specified a priori as a secondary outcome in the protocol and included data from all patients enrolled in the trial. The principal analysis was pre-specified as the incremental cost per QALY gained.

3.1 Reported Kaplan Meier estimates

Suppose we do not have access to individual data, now we need to estimate long term treatment effect of CRT+MT vs. MT. Cleland et al (2005) published their study in which they reported all cause mortality by Kaplan Meier estimate. Figure 2 was the Kaplan Meier estimated survival curve from their study.

3.2 Assessing the assumption of a constant hazard rate

Now we read a set of data points from the figure. Table 1 column 2 and column 3 were read from published Kaplan Meier curve.

Now we calculate the raw hazard ratio monthly based on the methods proposed in last section. For the last column in the table, the plot of monthly hazard rates against t was shown in figure 3. A trend curve was fitted into the scatter plot. This indicates that hazard rate is slightly decreasing with time but almost constant. Therefore the assumption of constant hazard rate (exponential distribution) will approximately fit the data. But Weibull distribution might provide a closer fit for this group of patients. First, we fitted exponential survival function as our baseline scenario and Weibull fitting in a sensitivity analysis.

- The estimated parameters by fitting Weibull function based on least square

regression methods were followings:

- Scale = 0.15 (95% CI 0.148 -0.152)
- Shape = 0.983 (95% CI 0.97 -1.00)
- The shape parameter is so close to 1. Indicates that Exponential could fit the curve well

- Exponential parameter =0.154 (95% CI 0.153 - 0.155)

Once the survival function for the control group was determined, the survival function for the treatment group can be deducted by multiplying the reported hazard ratio into the functions. Figure 4 shows the fitted curves and the observed curves from Kaplan Meier estimated.

3.3 Methods based on individual data

When individual data available from a clinical trial, we can a fit parametric survival model and use maximum likelihood methods to estimate the parameters and conduct goodness-of-fit test to select the best model. This is a standard method: most statistical software has the functionality of fitting parametric models. In order to validate the robustness of method used in last section, in which parameters were derived from published Kaplan Meier curve; in this section, we employed individual data to select the best model to fit the data. For all analysis in this section, we used SAS software.

First we fitted four parametric models into our data, Exponential, Weibull, Log logistic and lognormal. The Cox –Snell residual is used for initial check of model fitting, Figure 5 shows the Cox-Snell residuals from the four candidate models. We can see that both exponential and Weibull would provide a better fit than log logistic or lognormal.

Goodness of Fit Tests with Likelihood Ratio Statistic

Likelihood ratio statistic could be used to compare nested models. In this case, Exponential model is a special case of the Weibull model in which the shape parameter is restricted to be 1. Lognormal, Weibull and exponential are all nested within the generalized gamma model.

Table 1 reports negative 2 log likelihood statistics, to compare nested model, the difference between the log likelihood statistics is follow chi-squared distribution. Table 2 reported the p-value for different comparison.

However when we compare log logistic or lognormal with Weibull or exponential, they are not nested. Therefore we could not use likelihood ratio statistic, Akaike's Information Criterion (AIC) could use here.

AIC could be used to compare different parametric models by a statistic that trades off a model's likelihood against its complexity. It usually used to choose among non-nested models. A lower value of AIC indicates a better model.

$$AIC = -2LL + 2 * (c + a)$$

Table 3 and table 4 has indicated that Exponential would be the best fitting for our data. Table 5 reported the estimated hazard rate for MT group and the hazard ratios of CRT+MT compared to MT. The point estimation based on individual data is almost identical to the estimate based on Kaplan Meier curve alone. However the confidence

interval from the Kaplan Meier estimates is much narrow than the one from individual data. This is because, when estimating from Kaplan Meier curve, we have not considered individual variation. Ideally if we have reported confidence intervals around Kaplan Meier curve, we should construct confidence intervals around parameter estimates. However, this process is far from straight forwards. We need to inflate the variance of the estimated parameters based purely from a Kaplan Meier curve. In the CARE-HF case, inflating by 30 % gives reasonable results.

4 Discussion

On the whole, there are three important steps in the development of a useful model when survival time is a key input. First, identify the most relevant trial from the literature. Second, identify a probability distribution to represent the survival process. When data are available, this step typically begins by developing a hazard plot. Based on Kaplan-Meier estimate of survival function, several parameter models can be proposed, using graphical methods to investigate the suitability and estimated the parameters. Once a family of distribution is chosen, long-term survival can be obtained. Third evaluate the chosen distribution and the associated parameters for goodness of fit. Goodness of fit may be evaluated informally via graphical methods, or formally via statistical tests. The Chi-square and Kolmogorov-Smirnov tests are standard goodness-of-fit tests procedure.

An estimate of uncertainty is crucial for economic analysis, where it is quite inappropriate to take the observed survival function and assume that it has been measured without error. Variance estimation should account for both variation along the Kaplan Meier curves and variation in the reported Kaplan Meier estimates.

Reference:

1) John G.F. Cleland, M.D., Jean-Claude Daubert, M.D., Erland Erdmann, M.D., Nick Freemantle, Ph.D., Daniel Gras, M.D., Lukas Kappenberger, M.D., and Luigi Tavazzi, M.D., for the Cardiac Resynchronization — Heart Failure (CARE-HF) Study Investigators. The Effect of Cardiac Resynchronization on Morbidity and Mortality in Heart Failure. New England Journal of Medicine 2005; 352: 1539-49.

2) Gelber RD, Goldhirsch A, Cole BF, for the IBCSG. Parametric extrapolation of survival estimates with applications to quality of life evaluation of treatments 1993; 14:485-489 Controll Clin Trials

3) Glasziou PP, Cole BF, Gelber RD, Hilden J, Simes RJ. Quality adjusted survival analysis with repeated quality of life measures. 1998; 17:1215-1229.Statatics in Medicine

4) Cole BF, Gelber RD, Anderson KM. Parametric approaches to quality adjusted survival analysis.1994; 50:621-630. Biometrics

5) Alessandra Nardi, Michael. Comparing COX and parametric models in clinical studies. 2003;22:3597-3610 Statistics in Medecine

6) Elisa T. Lee and Oscar T. Go. Survival analysis in public health research. 1997, 18:105-34.Annual Review Public Health

7) Gerdtham UG, Zethraeus N. Predicting survival in cost-effectiveness analyses based on clinical trials. 2003 Summer;19(3):507-12.Int J Technol Assess Health Care

8) Jain AK, Strawderman RL. Flexible hazard regression modelling for medical cost data..2002 Mar;3(1):101-18.Biostatistics.

9) N. Neymark, I. Adriaenssen, T. Gorlia, S. Caleo, M. Bolla. Estimating survival gain for economic evaluations with survival time as principal endpoint: A cost-effectiveness analysis of adding early hormonal therapy to radiotherapy in patients with locally advanced prostate cancer. Health Economics, Volume 11, Issue 3, Date: April 2002, Pages: 233-248

10) Ruth Etzioni, Scott D. Ramsey, Kristin Berry, Martin Brown, The impact of including future medical care costs when estimating the costs attributable to a disease: a colorectal cancer case study. Health Economics, Volume 10, Issue 3, Date: April 2001, Pages: 245-256

11) Cohen A, Barnett O. Assessing goodness-of-fit of parametric regression models for lifetime data-graphical methods. Stat Med. 1995 Aug 30;14(16):1785-95

12) Mould RF, Lederman M, Tai P, Wong JK.. Methodology to predict long-term cancer survival from short-term data using Tobacco Cancer Risk and Absolute Cancer Cure models. Physical Mecicine Biology. 2002 Nov 21; 47(22):3893-924

13) Parmar, Mahesh K. B.. - Survival analysis : a practical approach. Chichester *Wiley.* 1995

14) Cox, D. R., David Roxbee, Analysis of survival data, London : Chapman and Hall, 1984. (Monographs on statistics and applied probability ; 21).

15) Collett, D. Modelling survival data in medical research. - London : Chapman & Hall, 1994

Table 1 Kaplan Meier estimates of survival function for time to all cause mortality

Months	Survival	Ln(t)	Ln(-Ln(S))	Ln(S)	Hazard
0	1.000				
1	0.983	0.33647	-4.04319	-0.01754	0.017
2	0.968	0.90287	-3.41587	-0.03285	0.015
3	0.963	1.19392	-3.26993	-0.03801	0.005
4	0.958	1.36949	-3.14151	-0.04322	0.005
5	0.945	1.68021	-2.87578	-0.05637	0.013
6	0.938	1.83524	-2.74291	-0.06438	0.008
7	0.930	2.00148	-2.62474	-0.07246	0.008
8	0.918	2.12026	-2.45223	-0.08610	0.014
9	0.895	2.24778	-2.19622	-0.11122	0.025
10	0.885	2.33537	-2.09854	-0.12264	0.011
11	0.880	2.43069	-2.05261	-0.12840	0.006
12	0.869	2.50689	-1.96592	-0.14003	0.012
13	0.854	2.59525	-1.84689	-0.15773	0.018
14	0.844	2.66491	-1.77354	-0.16973	0.012
15	0.839	2.72349	-1.73846	-0.17579	0.006
16	0.826	2.78501	-1.65468	-0.19115	0.015
17	0.816	2.83321	-1.59153	-0.20361	0.012
18	0.808	2.91416	-1.54560	-0.21318	0.010
19	0.800	2.96699	-1.49978	-0.22318	0.010

Table 2 Likelihood Ratio Statistic by Different Comparisons Between Nested Models

Likelihood ratio chi-square statistic	Contrast	Pr > ChiSq
0.36686	Weibull vs exponential	0.54472
2.10038	Gamma vs. exponential	0.3498
1.73352	Gamma vs. Weibull	0.18796
19.4868	Gamma vs. Lognormal	>0.000

Table 3 The Likelihood Statistic and AIC based on different parametric models

Model Fitting Information				
List	Distributions	No. Parameters	AIC	N2loglikelihood
1	Exponential	1	1309.03	1303.03
2	Weibull	2	1310.66	1302.66
3	Gamma	3	1310.93	1300.93
4	LLogistic	2	1313.63	1305.63
5	Lognormal	2	1328.42	1320.42

Table 4, estimated hazard ratio based on exponential model

Treatment class	Hazard yearly	95 % CI lower	96 % CI Upper	Mean Survival time in years	Life year saved
CRT + MT group	0.151	0.126	0.181	6.6	
Hazard ratio					
CRT+MT vs. MT	0.641	0.484	0.849		
MT group	0.097	0.081	0.116	10.3	3.7

Figure 1, Hazard plot against time from four cases studies

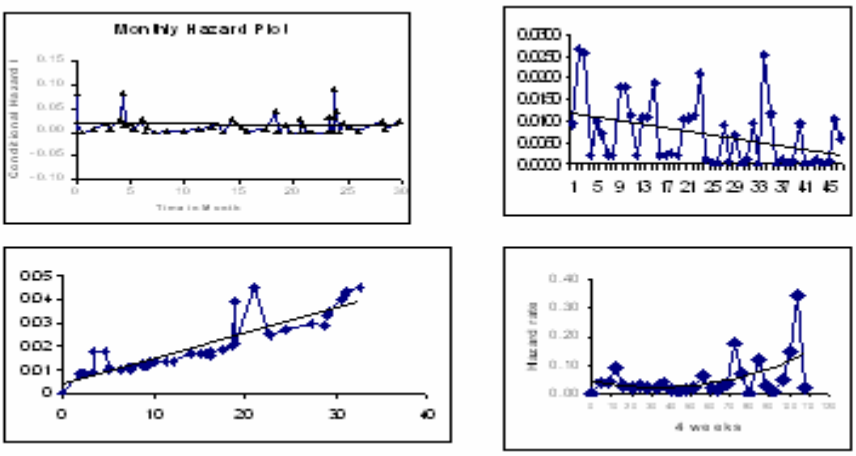

Figure 2 Kaplan–Meier Estimates of the Time to All Cause

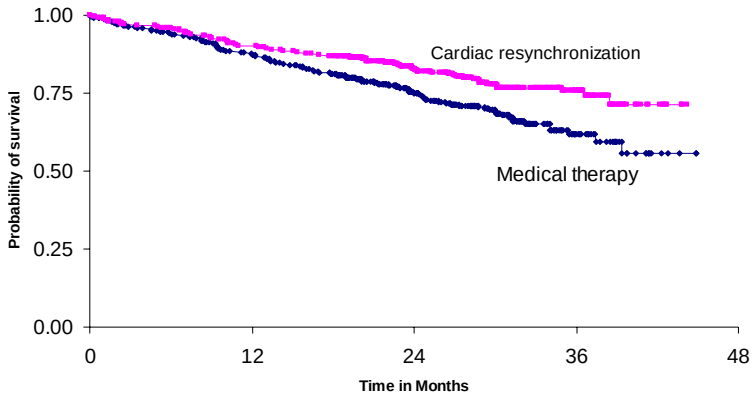

Figure 3: Hazard rates against time plot

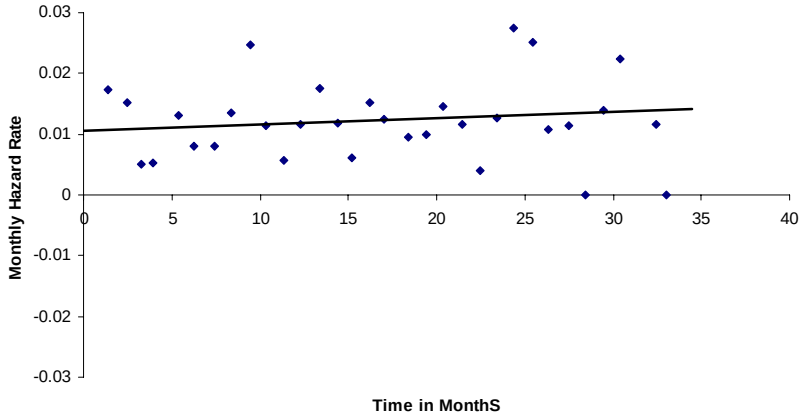

Figure 4 Estimated based on fitted curves and observed curves

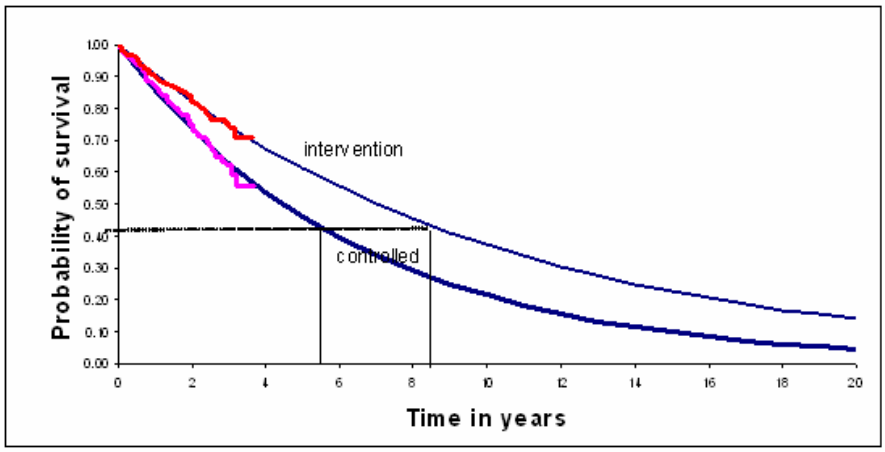

Figure 5 Cox-Snell residual plots by different survival functions