



# LUND UNIVERSITY

## Predicting survival in cost-effectiveness analyses based on clinical trials.

Gerdtham, Ulf; Zethraeus, Niklas

*Published in:*  
International Journal of Technology Assessment in Health Care

2003

[Link to publication](#)

*Citation for published version (APA):*  
Gerdtham, U., & Zethraeus, N. (2003). Predicting survival in cost-effectiveness analyses based on clinical trials. *International Journal of Technology Assessment in Health Care*, 19(3), 507-512.  
<http://swopec.hhs.se/hastef/papers/hastef0442.pdf>

*Total number of authors:*  
2

### General rights

Unless other specific re-use rights are stated the following general rights apply:  
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# PREDICTING SURVIVAL IN COST-EFFECTIVENESS ANALYSES BASED ON CLINICAL TRIALS

**Ulf-G. Gerdtham**

*Malmö University Hospital*

**Niklas Zethraeus**

*Stockholm School of Economics*

## Abstract

This study deals with the question of how to model health effects after the cessation of a randomized controlled trial (RCT). By using clinical trial data on severe congestive heart failure patients, we illustrate how survival beyond the cessation of an RCT can be predicted based on parametric survival models. In the analysis, we compare predicted survival and the resulting incremental cost-effectiveness ratio (ICER) of different survival models with actual survival/ICER. Our main finding is that the results are sensitive to the choice of survival model and that an extensive sensitivity analysis in the CE analysis is required.

**Keywords:** Cost-effectiveness analysis, Modeling, Confidence intervals

Several economic evaluations of new pharmaceuticals are being conducted alongside clinical trials where individual patient cost and health effect data are available (3;7). An advantage of using a controlled clinical trial as the base for the economic evaluation is that the results from the clinical study are of high internal validity in showing whether a new therapy has an effect. That the clinical study may be of relatively low external validity is a drawback, however, that is, it may not reflect the costs and health effects for patients in routine care, when the drug is out on the market.

A common and a very important methodological question in economic evaluation studies based on clinical trials is how to predict health effects after the cessation of the clinical trial. If the economic evaluation is based on a clinical survival study, the goal of modeling, thus, is to obtain an accurate estimate of the survival gain after the clinical study to be used in the economic evaluation. In general, the modeled gain in survival after the clinical trial is much larger than the observed gain stated; therefore, the manner of modeling can have major consequences for the cost-effectiveness (CE) of assessed therapies (1).

Analyzing the effect of a new treatment program in health economic studies, one would ideally like to carry out a study where patients are randomly allocated to “treatment” or

We thank John Kjekshus for providing us with data from the consensus and a ten-year follow-up study. Comments from Magnus Johannesson on a previous version of the article are also highly appreciated.

“no treatment” alternatives to study the causal effect of the program on survival in clinical practice. Usually, such data are not available at the time of the implementation of the new technology; however, the only available data for a newly registered drug often comes from a randomized controlled trial (RCT). In such trials, patients are commonly followed-up for a limited period of time, which suggests that survival must be modeled after the clinical study has been carried out. Such modeling can be made either on the basis of information within the clinical study or external information. In Jönsson et al. (3;4), external information is used and patients alive at the end of the follow-up are assumed to live according to actuarial data from national statistics. In these studies, the assumed expected survival time of ten years is reached by adding a constant survival time of ten years for all patients alive at the end of the clinical trial. To account for the fact that health effects will be discounted, this strategy is fulfilled by assuming that 5 percent of the surviving patients after the clinical study die in each year over a period of twenty years (3;4).

Methods for predicting survival based on observed patient data in a clinical trial can be characterized as parametric, semiparametric, and nonparametric. In Jönsson et al. (4), for example, a Weibull model was used in a sensitivity analysis to predict the expected conditional survival, whereas Raikou et al. (7) used a simulation model. However, there are several models that can be used for predicting survival, and a priori, there is no reason to prefer one model to the other. To evaluate the accuracy of the model, predictions would ideally be compared with the actual survival. By definition, this is not possible, however, because the lack of survival data beyond the cessation of the clinical trial is the reason for modeling survival.

The aim of this study is to illustrate how survival beyond the cessation of a clinical trial can be predicted based on different parametric survival models. In the analysis, we discriminate between the different survival models by the commonly used likelihood ratio (LR) test. We also compare the predicted survival/incremental cost-effectiveness ratio (ICER) based on the different models with the actual survival/ICER. All predictions are based on observed patient information contained within a clinical trial.

## DATA AND METHODS

Assume that a CE study is carried out alongside a clinical study, where a new therapy is compared with an existing standard therapy. Costs and health effects (measured as survival) are collected from a clinical trial with a mean follow-up time equal to  $X$ . It is assumed that all patients convert to the new therapy after the clinical study and that the new therapy, thus, is assessed against the old for a mean follow-up equal to  $X$ . The clinical trial showed a positive and significant effect on mortality, which means that there will be a gain in life-years within the follow-up time of the clinical trial. At the end of the clinical trial, more patients are alive in the active treatment group compared with the placebo group and the question is whether there will be further gains in survival after the clinical trial, when all patients convert to active treatment. How can information within the clinical trial be used to predict the expected remaining survival time for patients who are alive at the end of the follow-up? This question is investigated by using the below data.

### Data

The data are based on the consensus trial (9) and a follow-up study (8). The initial clinical study (9) compared enalapril with standard therapy in the treatment of congestive heart failure patients. The mean follow-up time in the clinical study was 0.515 years, and after the completion of the study, all patients were offered enalapril therapy (8). In the placebo group, 67 percent started to take enalapril, whereas 88 percent continued with it in the enalapril group. Individual patient data on life-years and whether they were alive at the end

of the follow-up were available in the clinical study. At the end of the clinical study, 77 of 127 patients were alive in the enalapril group compared with 58 of 126 in the placebo group. No comprehensive cost data were collected in the clinical study. To be able to calculate ICERs, individual cost data are randomly selected from a study comparing bisoprolol with standard therapy in heart failure (1). Individual cost and health effect data, thus, are available within the follow-up time in the clinical study. The health economic question is whether it is good value for money to implement the enalapril therapy (treatment 1) in addition to the standard therapy instead of using the standard therapy (treatment 0).

In this study, different modeling assumptions are made for patients who are alive at the end of the follow-up. The outcome of the modeling is then compared with the actual outcome presented in a ten-year follow-up study (8). At the ten-year follow-up, five patients were still alive in the enalapril treatment group, whereas one patient had been lost to follow-up in the placebo group. In our calculations, we assume that the follow-up study contains information of total survival after the end of the clinical study and those patients still alive at the 10-year follow-up die immediately after that point in time. The conditional mean survival time for patients alive at the end of the follow-up in the clinical study was 941 and 774 days in the enalapril and placebo group, respectively. This difference can possibly be explained by the fact that a higher share of the surviving patients in the enalapril group used enalapril also after the cessation of the clinical trial.

## Modeling

Our data on survival time have some characteristics that are important in selecting an estimation method. One is that the survival time distribution is usually skewed in some way, which violates the ordinary least squares assumption of normally distributed error terms. By definition, the survival time is also positive, whereas a normally distributed variable can take on both positive and negative values. Another characteristic is that a certain proportion of individuals has not reached the end-point of interest, that is, some individuals are still alive at the cessation of the clinical study, which means that such individuals are right censored. This finding calls for the application of duration data models incorporating the above characteristics (2). The random variable  $T$  is assumed to have a density function  $f(t)$ , reflecting the probability of the survival time being of length  $t$ , and a distribution function  $F(t) = \int_0^t f(s)ds = \Pr(T \leq t)$  defining the survival function  $S(t) = \Pr(T \geq t) = 1 - F(t)$ . The survival function shows the probability that the individual survives for at least  $t$  periods. From the survival function, one can define the hazard function  $\lambda(t) = \frac{f(t)}{S(t)}$ , which shows the mortality rate at time  $t$  conditional on surviving to time  $t$  (see Kiefer (5) and Lancaster (6) for surveys of duration models).

In the Results section, we estimate four common parametric survival models, that is, exponential, Weibull, lognormal, and generalized gamma models (2). These models are distributions for a non-negative random variable, with hazard functions displaying different kinds of behavior; for example, the hazard function for the exponential distribution is constant, whereas the hazards for the Weibull distribution are monotonically increasing or decreasing, depending on the shape parameter,  $p$  (2). The hazard function of the generalized gamma model is very flexible, allowing for a large number of shapes and the exponential, Weibull, and lognormal models are special cases of the generalized gamma model. Thus, these models can be tested as null hypotheses against the alternative generalized gamma model by the use of a LR  $\chi^2 \sim$  test.

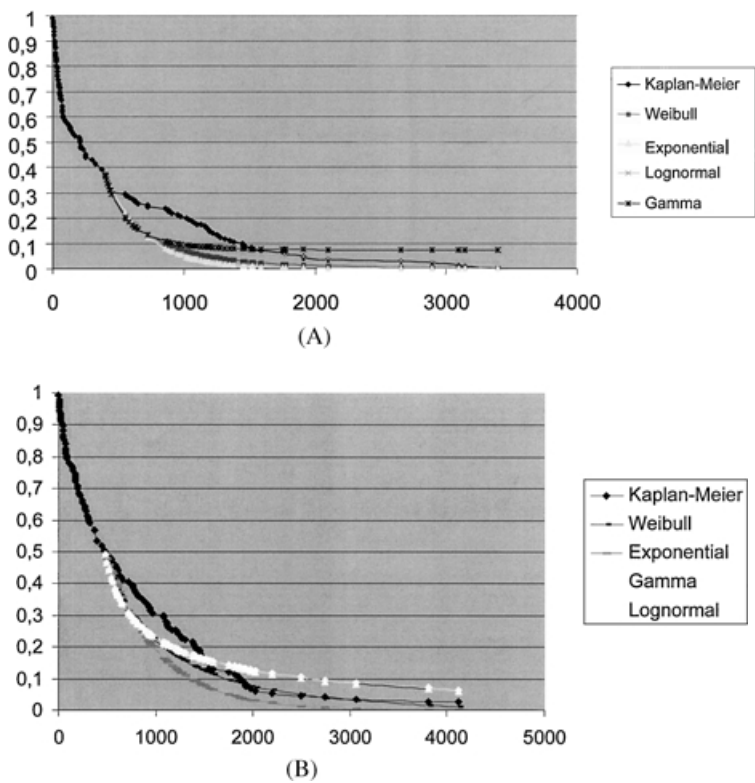
## RESULTS

Table 1 shows the actual and expected survival (discounted and undiscounted) and costs with and without treatment based on different model alternatives. According to the LR  $\chi^2 \sim$  test,

**Table 1.** Mean Survival Time (Days) and Costs within the Clinical Trial and Total Actual and Modeled Mean Survival, Respectively

	Undiscounted				Discounted			
	Standard	Enalapril	Difference	ICER	Standard	Enalapril	Difference	ICER
Costs (SEK)	21,361	34,711	13,350		21,361	34,711	13,350	
Health effects								
<i>True survival</i>								
Within trial	159.37	215.32	55.95		159.37	215.323	55.950	
Total follow-up	515.53	785.95	270.42	18,019	502.26	767.280	265.01	18,387
<i>Modeled Survival</i>								
Gamma	927.84	1227.50	299.65	16,261	865.87	1145.88	280.01	17,402
Lognormal	902.74	1194.44	291.69	16,705	844.03	1117.11	273.07	17,844
Weibull	574.23	761.75	187.51	25,985	555.63	737.25	181.61	26,830
Exponential	410.64	546.27	135.63	35,925	405.11	539.00	133.88	36,395

ICER, incremental cost-effectiveness ratio.



**Figure 1.** Actual survival (Kaplan-Meier) compared with modeled survival for the placebo group (A) and for the enalapril group (B).

the Weibull and the exponential model but not the lognormal model are rejected against the gamma model at the 1 percent level of significance. However, the model best predicting actual survival (discounted or undiscounted) in the enalapril and placebo group is the Weibull model. The exponential model systematically underestimates survival, whereas the gamma and lognormal models overestimate survival. The models best predicting the difference in

survival are the lognormal and gamma ones, because these overestimate the survival in the treatment and the placebo group to approximately the same extent. The Weibull model, on the other hand, overestimates survival in the placebo group and underestimates survival in the enalapril group. Thus, when comparing the ICERs (discounted or undiscounted) derived from the models with the “true” ones, the lognormal and gamma models are the most accurate (Table 1).

Note that after the clinical study, we assume equal expected survival in the two treatment groups. However, the actual observed data show that the conditional expected survival time is higher in the enalapril group (941) compared with the placebo group (774). Thus, even if we predicted the correct conditional expected survival in the enalapril group and used that for the placebo group, the true difference in survival would be understated. Models “correctly” overestimating conditional expected survival in the enalapril group give survival differences that equal the true difference.

In Figure 1, the survival curves based on actual data and models are presented for the enalapril and placebo groups. Note that the survival curve based on the observed data is above the modeled survival curves at the beginning of the prediction period (for the first 1,500 days) and, thereafter, comes close the modeled survival curves.

## SUMMARY AND CONCLUSION

This study investigates the question of predicting survival in CE studies based on the information contained in an RCT with a given follow-up. In the study, we predict survival by using different parametric survival models (generalized gamma, lognormal, Weibull, and exponential), which are tested against each other by using the likelihood ratio test. Furthermore, the model predictions are compared with observed true survival.

The Weibull model most accurately predicts actual survival in the enalapril group. However, this finding was not confirmed by the statistical tests. The used likelihood ratio test rejected the Weibull (and the exponential) model but not the lognormal model against the gamma model. Thus, for this particular data, the gamma and lognormal models would be selected based on the statistical tests, although the Weibull model generated the most accurate survival predictions. The lognormal and gamma models most accurately predict the observed difference in survival between the two therapies. The reason for this finding is that these models overstate the conditional survival in the enalapril group that reduces the underestimation of the actual difference in survival resulting from the Weibull model. A model exactly predicting the conditional actual survival in the enalapril group will underestimate the true difference in survival.

Our conclusion is that statistical tests discriminating between models used to predict survival outside the sample should be complemented by an extensive sensitivity analysis, because it is not obvious that the model performing best in the statistical tests generates the best survival predictions.

To assess our model predictions, we ideally need health effect data from an extended RCT, that is, a study that also continues after the cessation of the initial RCT and where all patients use enalapril in a controlled setting. Our data are based on an open setting where the two patient groups are offered the new therapy. In what way do our data differ from ideal data? The data in this study reflect actual clinical practice (effectiveness) and not a controlled situation (efficacy). After the cessation of the clinical trial, a lower share of patients in the initial placebo group that begins with enalapril compared with the share of patients continuing with enalapril in the enalapril group. In a controlled setting, no difference in the share of patients in active therapy is expected. However, the patients in this study are in severe disease states and probably have frequent contacts with health care, which means that clinical practice is rather close to the controlled environment.

Instead of using an RCT as a base for the economic evaluation, a controlled trial conducted under more naturalistic circumstances can be used for the health economic evaluation. Such a study would be characterized by both high internal and external validity. However, at the time of the registration of a new chemical entity, the “only” available information is usually that offered by the RCT. The appropriateness of using an RCT should be assessed and discussed and can vary, depending on the patient group under study.

## REFERENCES

1. Ekman M, Zethraeus N, Jönsson B. Cost-effectiveness of bisoprolol in the treatment of chronic congestive heart failure in Sweden. Analysis using data from the cardiac insufficiency bisoprolol study II. *Pharmacoeconomics* 2001;19:1-17.
2. Greene WH. *Econometric analysis*. New York: Prentice Hall International; 2000.
3. Jönsson B, Cook JR, Pederswen TR. The cost-effectiveness of lipid lowering in patients with diabetes: Results from the 4S trial. *Diabetologia*. 1999;42:1293-1301.
4. Jönsson B, Johannesson M, Kjekshus J, et al. Cost-effectiveness of cholesterol lowering. Results from the Scandinavian Simvastatin Survival Study (4S). *Eur Heart J*. 1996;17:1001-1007.
5. Kiefer NM. Economic duration data and hazard functions. *J Econ Lit*. 1988;26:646-679.
6. Lancaster T. *The econometric analysis of transition data*. Cambridge: Cambridge University Press; 1990.
7. Raikou K, Gray A, Briggs A, et al. for the UK prospective diabetes study group. Cost effectiveness analysis of improved blood pressurecontrol in hypertensive patients with type 2 diabetes: UKPDS 40. *BMJ*. 1998;317:720-726.
8. Swedberg K, Kjekshus J, Snapinn S, for the consensus investigators. Long term survival in severe heart failure patients treated with enalapril. *Eur Heart J*. 1999;20:136-139.
9. The Consensus trial study group. Effects of enalapril on mortality in severe congestive heart failure. Results of the cooperative north Scandinavian enalapril survival study (consensus). *N Engl J Med*. 1987;316:1429-1435.