

The Mean, Median, and Confidence Intervals of the Kaplan-Meier Survival Estimate—Computations and Applications

Chris BARKER

This short note points out estimators of the mean, median, and the associated confidence intervals of the Kaplan-Meier product limit estimate. Some uses of the estimator of the mean are described. In addition, differences among popular software packages in the calculation of both the mean and median and associated confidence intervals are demonstrated and are due to default settings in the software. Simple examples of the calculations are provided using S-Plus, R, SAS, Stata, and SPSS.

KEY WORDS: Median Survival; Pharmacoeconomics; Restricted Mean; Survival Analysis Summary Statistics.

1. BACKGROUND

This article was motivated by the author's involvement in a Pharmacoeconomics project to develop a cost-effectiveness model. Preparation of the economic model required an estimate of the mean of a survival distribution. At the time, the project economist provided a mean of a Kaplan-Meier (KM) computed in SAS, and the author provided a mean computed in S-Plus. The output and interpretation of the cost effectiveness model changed dramatically due to differences in the two means. On investigation it became clear that the algorithm for the calculation of the mean in SAS differed from S-Plus. The values differed dramatically, because a single patient had the longest censored survival. The reason and an example of the difference in means will be described later. In assembling the calculations of the mean in several software packages, differences were also noted in the lower and upper confidence limits for the median. As will be discussed later, these differences in the CIs are due to differences in the method for computing the confidence intervals of the survival distribution.

2. ESTIMATION OF THE MEAN

The median is commonly used to summarize the Kaplan-Meier Survival Estimate (Kaplan and Meier 1958). The mean of the KM Survival Estimate is less frequently used as a summary statistic. In contrast, health economic summaries of

cost effectiveness may involve survival, and the mean is typically the preferred summary statistic.

Miller (1981) outlines three ways to compute the mean of the KM:

- (1) Redefinition of the last observation, essentially an area under the curve including the longest observation, if that is censored. For the calculation, the longest survival time is treated as an "event,"
- (2) Irwin's restricted mean (Irwin 1949), later described by Kaplan and Meier (1958) as $\hat{\theta} = \int_0^{T_k} \hat{S}(t) dt$, where T_k is a time point for computing the mean.
- (3) Susarla and Ryzin's variable upper limit (Susarla and Ryzin 1980). For this calculation denoting $\hat{\theta} = \int_0^{s_n} \hat{S}(t) dt$, where $\{s_n\}$ is a sequence of numbers converging to ∞ .

Klein and Moeschberger (1997) cite two additional methods that could be applied for estimating the mean:

- (1) Efron's suggestion to estimate $S(t)$ by 0 beyond the largest study time t_{\max} and
- (2) Gill's suggestion to estimate $S(t)$ by $S(t_{\max})$ for $t > t_{\max}$, which corresponds to assuming that individual would die at ∞ .

The calculation of the Kaplan-Meier mean in SAS Proc Lifetest (SAS Institute Inc. 2002–2004), which computes the area under the curve up to the last failure time, does not appear to have been described in the literature. Proc Lifetest reports in the output that the mean is an underestimate when the longest survival time is censored. SAS V9 also provides an option to restrict the calculation of the mean to a specific time. Stata provides an option to compute the mean using an extrapolation of the survival distribution described in Brown, Hollander, and Korwar (1974).

3. EXAMPLE

Kaplan-Meier estimates and summary statistics were prepared using the following fictitious survival time data, with the longest observation censored, where + denotes censoring, (10, 15, 23, 30, 35, 52, 100+). The mean and median and its confidence intervals are displayed in Table 1. The variance of the mean is based on the Greenwood (1926) estimator of the variance of the survival distribution.

4. CONFIDENCE INTERVAL FOR THE MEDIAN

The upper and lower limit of the confidence intervals for the median in Table 1 differ among the packages considered here. The main reason is because different methods are used for determining the confidence interval for the survival function.

Klein and Moeschberger (1997) outline several methods for calculating the confidence interval for the survival function:

Chris Barker, Ph.D. is a statistical consultant to the pharmaceutical industry in the San Francisco Bay Area (E-mail: chrismbarker@yahoo.com). The author appreciates the help of several colleagues in preparing some calculations and answering questions about the software packages, Ted Karrison, (U. Chicago), Andy Bogart (U. Washington), Jon Peck (SPSS Inc.), Terry Therneau (Mayo Clinic), Paul Savarese (SAS Institute). The author is responsible for any errors in the manuscript.

Table 1.

	Procedure	Confidence Interval Method	Median (95% CI)	Mean (SE)
S-Plus 7	Survfit(Surv())	Log	30 (15, NA)	37.9 (10.7)
R (2.8.0)	Survfit(Surv())	Log	30 (15, inf)	37.9 (10.7)
SAS (Version 9.2)	Lifetest	Linear	30 (15, 52)	31 (6.4)
		Log	30 (15, <missing>)	31 (6.4)
		Log-log	30 (10, 52)	31 (6.4)
SAS (Version 9.1)	Lifetest	Linear	30 (15, 52)	31 (6.4)
		Log	30 (15, 52)	31 (6.4)
		Log	30 (10, 52)	31 (6.4)
		Log	30 (12.0, 47.9)	37.9 (10.7)
SPSS (Version 16)	Survival module	Log-log	30 (10, 52)	37.9 (10.7)
Stata (Version 10)	Survival module	Log-log	30 (10, 52)	37.9 (10.7)
	Extended mean			45.2

NOTE: Stata offers a method to impute the tail of the distribution and compute an “Extended” mean.

- (1) Linear,
- (2) Log-transformed,
- (3) Arcsine-square root.

Figure 1 presents the estimated survival curve for the fictitious data, computed in SAS V9 using three different methods for estimating the confidence intervals. Figure 2 presents the survival curve and the default estimate from S-Plus. It may be noted that the confidence limits in S-Plus extend to the longest survival time. The default SAS Proc Lifetest output delivery system (ODS) output does not extend beyond the last failure time.

S-Plus and R compute the confidence interval for the median using the method of Brookmeyer and Crowley (1982) based on inverting a sign test (Therneau, T., personal communication, September 23, 2008). The estimates of the lower and upper limit may be obtained by drawing a horizontal line at 0.5 on the vertical axis, and drawing a perpendicular line to the x-axis at the points where the horizontal line intersects the confidence intervals. “NA” or “inf” is reported where the horizontal line does not intersect a confidence interval.

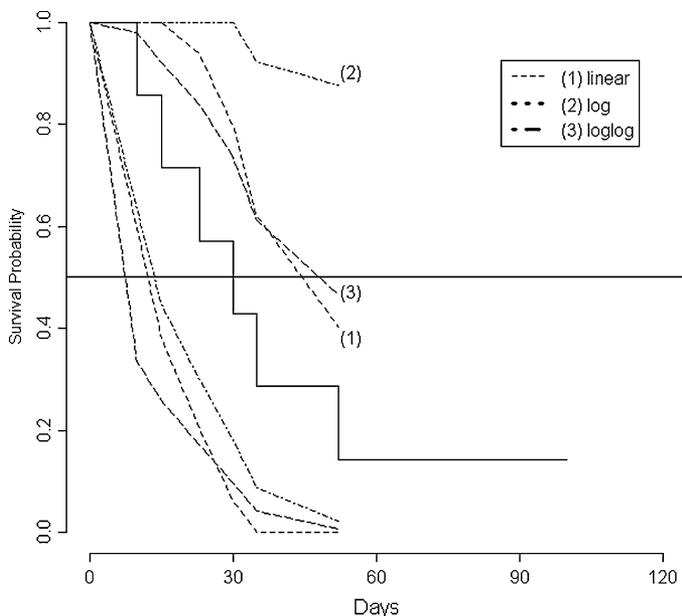


Figure 1. Kaplan-Meier Survival Estimate with three estimates from SAS Version 9.2 of the 95% confidence intervals.

The confidence intervals in SAS Proc Lifetest for the median (quartiles) are given by: $I_{0.50} = (t : (1 - \hat{S}(t) - 0.50)^2 \leq c_{\alpha} \hat{\sigma}^2(\hat{S}(t)))$ where c_{α} is the upper α percentile of a central chi-squared with 1 degree of freedom (SAS Institute Inc. 2008). SAS Version 9.2, uses a generalization of the Brookmeyer and Crowley (1982) sign test based on the type of the confidence interval (SAS Technical Support, personal communication, September 2008, and SAS online help 2002–2005). This differs from SAS Version 9.1 where the confidence intervals for the median, and other quartiles, are based only on the linear function.

SPSS (2008) estimates the standard error of the median as $se(P_i) = P_i \sqrt{\sum_{j=1}^i (q_j/r_j p_j)}$ where the proportion terminating is $q_i = d_i/n_i$. The confidence interval is based on a normal approximation, $1.96 * SE$.

5. APPLICATIONS OF THE RESTRICTED MEAN OF THE KAPLAN MEIER ESTIMATOR

In Health Economics, there is an important implementation of the restricted mean, in a “QTWIST” analysis. The methodology of Glasziou uses a “partitioned survival analysis” (Glasziou, Simes, and Gelber 1990). The method applies the restricted mean (RM) for the estimation of “time” spent in a particular state of health.

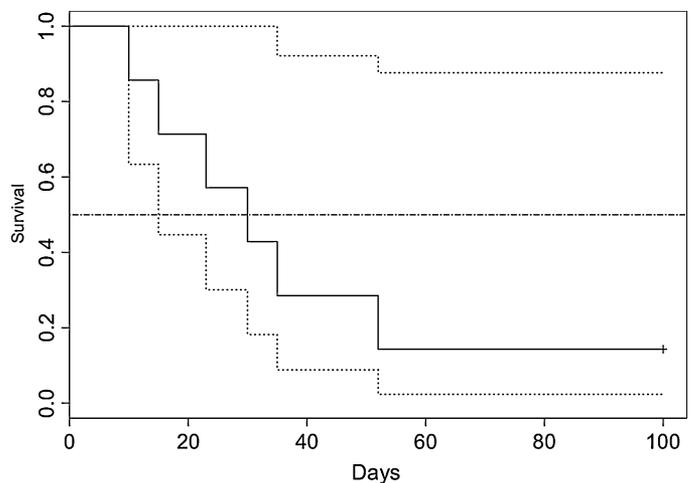


Figure 2. Survival Estimate with confidence limits from S-Plus 7.0

Glasziou et al.'s model assumes progressive health states, where each change of health state is a separate endpoint. Denoting each state by $-i$ -, in the following equation, then the restricted mean time in each state is computed by Equation (1), with the common restriction time, say L . Denoting the RM for this health state model as $\mu_i(L)$, then the estimate of duration of each state is $\mu_{i+1}(L) - \mu_i(L)$, and $\mu_0(L) = 0$.

Zhao and Tsiatis (1997) develop a related estimator for Quality Adjusted Lifetime, using a Martingale formulation. Karrison (1987) and Zucker (1998) have developed regression methods for the restricted mean that permit adjustment for covariates. Software for regression models for the restricted mean, or QTWIST are not currently available in the packages reviewed in this article.

6. CONCLUSIONS

Care must be taken to review the default settings in Kaplan Meier survival analysis software for computing the mean, the median, and the associated confidence intervals. The statistician should select the particular method of estimation of the mean for the Kaplan Meier estimate of survival, including prespecification of the time-point of restriction of the mean. The confidence limits for the median depend on the choice of the confidence intervals for the survival curve and this also should be prespecified for the statistical analysis.

REFERENCES

Brown, J., Hollander, M., and Kornwar, R. (1974), "NonParametric Tests of Independence for Censored Data with Applications to Heart Transplant Studies," in *Reliability and Biometry, Statistical Analysis of Lifelength*, eds. F. Proschan, R. Serfling, Philadelphia: SIAM, pp. 327–354.

Brookmeyer, R., and Crowley, J. (1982), "A Confidence Interval for the Median Survival Time," *Biometrics*, 38, 29–41.

Glasziou, P., Simes, R., and Gelber, R. (1990), "Quality Adjusted Survival Analysis," *Statistics in Medicine*, 9, 1259–1276.

Greenwood M. (1926), "The Natural Duration of Cancer," in *Reports on Public Health and Medical Subjects*, London: Her Majesty's Stationery Office," 33, pp. 1–26.

Irwin, J. O. (1949), "The Standard Error of an Estimate of Expectation of Life, with Special Reference to Expectation of Tumourless Life in Experiments with Mice," *The Journal of Hygiene*, 47 (2), 188.

Kaplan, E. L., and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53, 457–481.

Karrison, T. (1987), "Restricted Mean Life with Adjustment for Covariates," *Journal of the American Statistical Association*, 82 (400), 1169–1176.

Klein, J., and Moeschberger, M. (1997), *Survival Analysis—Techniques for Censored and Truncated Data*, New York., Springer.

Miller, R. (1981), *Survival Analysis* Wiley.

R Development Core Team (2008). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

R version 2.8.0 CRAN, New York, NY.

Susarla, Y., and van Ryzin, J. (1980), "Large Sample Theory for an Estimator of the Mean Survival Time from Censored Samples," *Annals of Statistics*, 8, 1002–1016.

SPSS Inc., SPSS Statistics 17.0 Algorithms, (2008).

SAS Institute Inc. (2004), *SAS/STAT 9.1 User's Guide*, Cary, NC. SAS Institute Inc.

———. (2008), *SAS/STAT® 9.2 User's Guide*, Cary, NC: SAS Institute Inc.

———. (2002–2005), *SAS Online Doc, 9.1.3*, Cary, NC: SAS Institute Inc.

S-Plus Version 7.0, TIBCO, formerly Insightful Corporation, Seattle WA.

Stata Version 10, Stata Corporation, College Station, TX.

Zhao, H., and Tsiatis, A. (1997), "A Consistent Estimator for the Distribution of Quality Adjusted Survival Time," *Biometrika*, 84, 339–348.

Zucker, D. (1998), "Restricted Mean Life with Covariates: Modification and Extension of a Useful Survival Analysis Method," *Journal of the American Statistical Association*, 93, 000–000.