Understanding survival analysis: actuarial life tables and the Kaplan–Meier plot

ABSTRACT

Survival analysis is a set of methods used to study the time between enrollment in a study and the occurrence of an event of interest. Two methods are commonly used: actuarial life tables and the Kaplan–Meier approach for survival analysis. A good understanding of both these methods is useful when reading and appraising the literature concerning prognostic and interventional studies. Kaplan–Meier curves are widely used as they enable analysis of incomplete sets of data (i.e. after patients withdraw from studies or are lost to follow up). This review explains these two methods and gives practical examples of their use.

Survival analysis is a statistical method of analysing the time to occurrence of an index event. Actuarial life tables and Kaplan–Meier curves are graphical means of representing survival analysis. A working knowledge of these methods is useful to help interpret the medical and scientific literature.

Important definitions

The index event (end point)

This is often survival or death, but other end points can be used (e.g. tumour relapse, renal transplantation in patients with end-stage renal failure) (Stel et al, 2005). Survival analysis is widely used in the medical literature to analyse aspects such as recovery rates, mortality rates and the effectiveness of treatment (Jemal et al, 2008).

The time to event

This is conventionally termed the survival time, even if the chosen event is not death. The event could be a single or multiple end points (Wanner et al, 2005). There are two common problems in calculating the mean survival time for a cohort of patients: different start times for given subjects (secular dates) and data loss over time (censoring).

Mr Ahmed Barakat, Clinical Fellow Trauma and Orthopaedic Surgery, Department of Trauma and Orthopaedics, Princess Royal Hospital, Haywards Heath, West Sussex RH16 4EX

Dr Aaina Mittal, Foundation Year Doctor, Department of Trauma and Orthopaedics, Princess Royal Hospital, Brighton and Sussex University Hospitals, Haywards Heath, West Sussex

Mr David Ricketts, Consultant Orthopaedic Surgeon, Department of Trauma and Orthopaedics, Princess Royal Hospital, Haywards Heath, West Sussex

Mr Benedict A Rogers, Consultant Orthopaedic Surgeon, Department of Trauma and Orthopaedics, Princess Royal Hospital, Haywards Heath, West Sussex

Correspondence to: Mr A Barakat (ahmedharoonbarakat@gmail.com)

Serial time

The different dates in the calendar at which subjects enter the study are known as the calendar dates or secular dates. To correct for different start times, all subjects in the cohort begin analysis at the same standardized point of time: time zero. The serial time for each subject is the time duration from which the subject started the study (time zero) until the occurrence of either the event or being censored.

Censoring

Censoring is loss of data during a study (as a result of participant withdrawal or being lost to follow up) or at the end of the study (because the event has not occurred and the subject survives until the end of the study period).

The distribution of censored subjects over time is important. For example, if there were many censored subjects a curve representing treatment effectiveness could indicate that subjects withdrew from the study to pursue different therapies (Rich et al, 2010).

Gradual omission of censored observations reduces the sample size and correspondingly causes reduction in the accuracy of survival estimates. Less than half of 319 published randomized controlled trials using Kaplan–Meier methods accurately included the details of patients lost to follow up, potentially interfering with the conclusions (Vervölgyi et al, 2011).

Two common methods allow for censored data: the actuarial life table and the Kaplan–Meier curve.

Time intervals

The main difference between these survival methods is that in actuarial life tables, study duration is expressed in equally spaced set intervals at which follow up occurs. Thus, survival time (i.e. time to event) is expressed in the number of intervals before the event or censoring. However, with the Kaplan–Meier approach, observed event times and censoring times are used to express survival times.

The conditional probability

The probability of surviving an interval without having the event.

The cumulative probability

The overall chance of survival to any time point.

Table 1. An example of the basic information required to construct a life table for a cohort of 100 subjects followed up for 4 years. The intervals refer to the preset follow-up periods. At each interval the number of patients entering the interval and remaining after events and censoring is calculated (i.e. number at risk at the start and at the end of interval). Conditional and cumulative probability of survival are derived for each interval

Interval (years)	No. at risk at start of the interval	Events	Censoring	No. at risk at end of the interval	Conditional probability of survival	Cumulative probability of survival	Incidence rate of events
1	100	20	0	80	0.8	0.8	0.2
2	80	15	20	45	0.75	0.6	0.25
3	45	10	15	20	0.66	0.4	0.33
4	20	10	0	10	0.50	0.2	0.50

The incidence rate

A measure of the number of events per unit of time and so a measure of rate rather than risk. It does not account for the decay of the population at risk (i.e. number at risk is gradually decreasing with the occurrence of each event) and should be differentiated from the cumulative probability.

The actuarial life table

This is a simple method of analysing survival. Survival probabilities are calculated at fixed intervals. Life table analysis is often presented graphically as a survival curve that plots time *vs* cumulative survival.

These graphs are constructed as follows (*Table 1*, *Figure 1*):

X-axis

This graphs time. The serial time starts at time zero for all subjects. The follow-up duration and status (i.e. has the event occurred or not) are set at predetermined follow-up intervals. The follow-up intervals divide the follow-up time of the cohort into a number of discrete pieces. In the above example (*Table 1*, *Figure 1*), a cohort of 100 subjects has been followed up for 4 years with preset yearly follow up (i.e. four intervals).

Y-axis

This graphs the cumulative probability. The curve will step (change in cumulative probability) at the preset intervals and not after each event occurrence as in Kaplan–Meier curves. In *Figure 1*, the curve steps down as the cumulative probability declines at the end of each interval. At the start of the study, the cumulative probability for survival was 1.0, and at the end of the first, second, third and fourth years it was 0.8, 0.6, 0.4 and 0.2 respectively.

Censoring

To allow inclusion of censored data life tables assume that withdrawal or censoring occurs in the middle of a studied interval (uniform timing). Censored patients are usually indicated by tick marks (or dots) along the interval in which they were censored. Censoring will not cause the cumulative probability to drop (i.e. the curve to step down or the interval to terminate), but it will decrease the number of patients at risk available for the next interval.

Conditional survival probability

Conditional survival probability is calculated for each interval as follows:

1 – [(number of events occurring in that interval)/(number of patients at risk at the start of the interval - censored subjects)]

For example, looking at *Table 1*, at the end of the first interval only 20 subjects had the event out of the 100 subjects with no censoring. The conditional probability of survival for the first interval is 0.8.

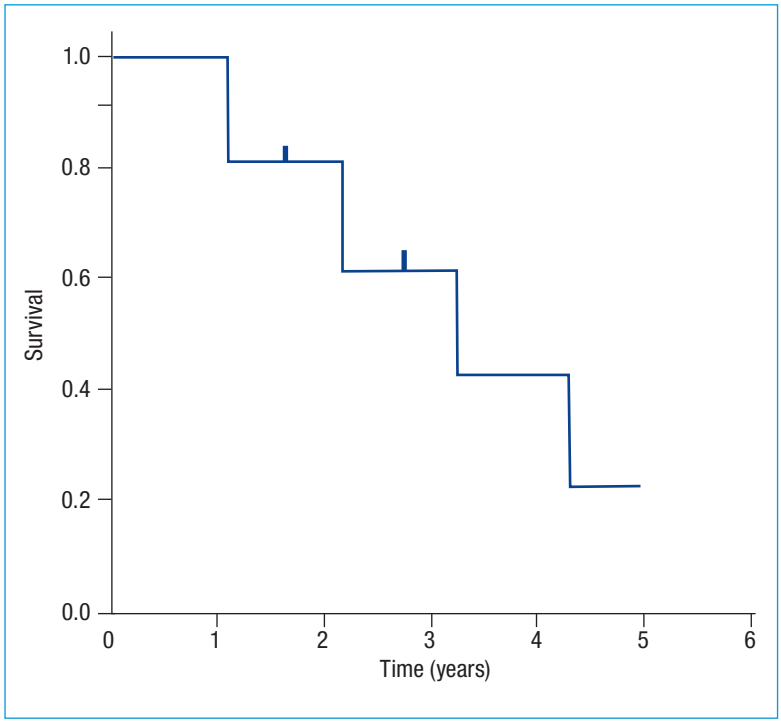

Figure 1. An actuarial survival curve derived from the data tabulated in *Table 1*. The X-axis shows the intervals. The Y-axis shows the cumulative survival probability. Vertical short lines pointing upwards refer to censoring. The survival rates are calculated at the follow-up intervals. The curve will step (change in cumulative probability) at the preset intervals.

Cumulative survival probability

Cumulative survival probability is the product of conditional probabilities of survival up to that interval. For example, the cumulative probability of survival at the end of the second interval would be 0.75 x 0.8 = 0.6.

Incidence rate

Incidence rate is calculated as [number of events occurring in that interval / (number of patients at risk at the start of the interval - censored subjects)]. The incidence rate of events at the end of the fourth interval is [10/(20-0)] = 0.5.

The usefulness of actuarial life tables is limited because analysis is at preset intervals. It neglects details of the study between two dates of calculation.

Table 2. A comparison of Kaplan–Meier survival curves *vs* actuarial tables

Actuarial life table	Kaplan–Meier survival curve
Allows for censored data and cumulative probability calculations	Allows for censored data and cumulative probability calculations
Time dependent (calculations made at fixed time intervals)	Event dependent (recalculated after each event)
Unreliable in small data sets	Reliable in small data sets
Uniformity assumption (uniform timing of censoring assumed)	No assumption is made about the timing of censoring

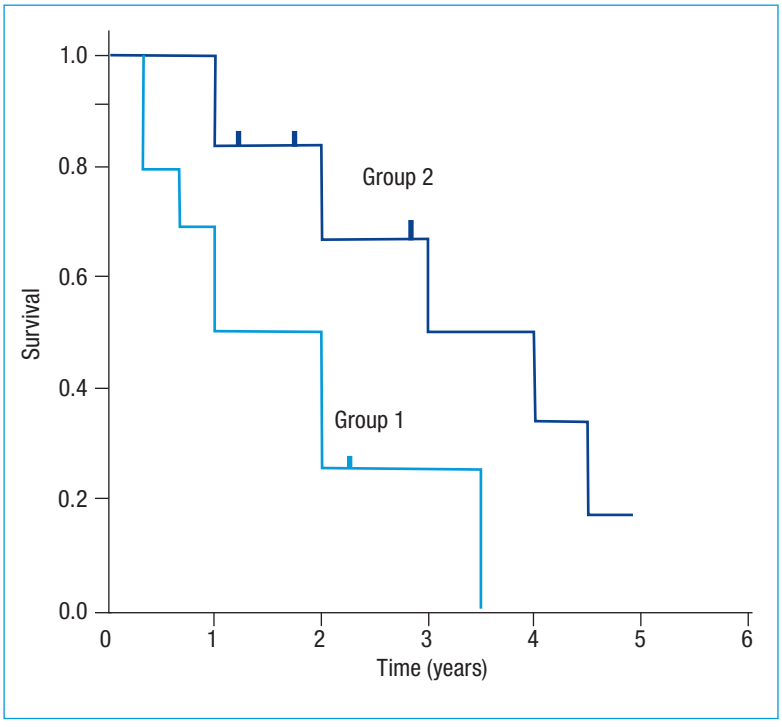

Figure 2. Kaplan–Meier survival curve for an intervention-control study. Each cohort consists of 10 patients. The X-axis shows the serial times (in years). The Y-axis shows the survival probability. Vertical short lines pointing upwards refer to censoring while vertical longer lines pointing down refer to occurrence of events.

The Kaplan–Meier curve

The Kaplan–Meier method recalculates the survival rate each time the event of interest occurs rather than at preset follow-up intervals as in actuarial life tables. Kaplan–Meier analysis is thus event dependent and not duration dependent. This provides more accurate measurement of survival rates at different intervals. With this method, rather than listing serial time in terms of predetermined intervals, time is marked in terms of events (Kaplan and Meier, 1958).

The relative advantages of actuarial tables and of Kaplan–Meier curves are shown in *Table 2*.

A Kaplan–Meier curve is constructed as follows (*Figure 2*):

X-axis

This graphs time. Again, it is assumed that all participants started the study at the same point of time (i.e. time zero).

Horizontal lines along the curve are the survival times or intervals and show the time starting from a defined point to the occurrence of the event (Akl et al, 2009). Steps in the graph correspond to the times at which events were observed, usually connected by vertical lines for cosmetic value (Vervölgyi et al, 2011).

It follows that the higher the number of participating subjects, the more the events and thus the more frequent the steps in the curve. Curves with large steps usually have a smaller number of subjects and correspondingly less accuracy.

The longer the observation period, the higher the accuracy.

Y-axis

This shows participants who have not had the event nor been censored.

The survival probability is 1 at the study's commencement (i.e. no events have yet occurred). At this point the entire cohort would be 'at risk' of subsequent events. Follow-up periods or serial times would be graphed on the X-axis and would start from time zero till the end of follow up.

As failures (events) begin to occur during the follow-up period, the survival probability decreases. The accuracy of the study is higher at the left of the curve (higher sample size) and reduces with time (Drüeke et al, 2006).

The slope of the curve is a function of the survival duration. The steeper the slope, the more frequent the event occurrence and thus the less the survival duration between events.

Censoring

This is indicated by short vertical lines or tick marks of equal lengths at the relevant serial time. These tick marks simply indicate the times at which censoring occurred but not the numbers of censored subjects. These subjects would not affect the survival probability or cause the curve to step down, but they would decrease the remaining number of subjects at risk.

To compensate for censoring two tails to the curve can be displayed. The first displays a standard survival curve of participants who have had the event. The second shows a worst-case curve which assumes that patients lost to follow up have reached the end point or the event of interest (Kaplan and Meier, 1958). This is demonstrated in *Figure 3* which shows that the true Kaplan–Meier estimate would actually lie between these two tails (i.e. the actual scenario curve and the worst-case scenario curve). The reader can then be certain that the true survival curve lies somewhere between the worst-case and standard survival curves (Murray et al, 1993).

Conditional survival probability

This is calculated as follows:

1 – [(number of events)/(number of patients at risk)].

Consider group 1 in *Figure 2*: after 4 months, two events occurred. The conditional survival probability would be 0.8 and the number at risk would decrease to 8. Another event occurred at 8 months leading to a conditional probability of 0.875 (1– [1/8]).

Cumulative survival probability

This is the product of conditional probabilities up to a time point. From the above example, the cumulative survival probability at 8 months would be 0.8 x 0.875 = 0.7.

The cumulative survival probability can be calculated for different groups and the groups then compared. The log-rank test for statistical significance or the Cox proportion hazard test are used to compare rates of having an event in one group against another (i.e. intervention *vs* control) (Marubini and Valsecchi, 1958).

Reliability

The number of participants remaining in the follow-up group is an important variable determining the reliability of the data. With numbers as low as 10 the possible error is approximately 20% (Peto et al, 1977). This can be corrected by adding 95% confidence intervals for the cumulative probabilities especially at the tail of the curve.

Median survival time

This can be estimated by drawing a vertical line from where the curve crosses the 50th percentile down to the X-axis. In *Figure 2*, the median survival time for group 1 is 1 year while for group 2 it is 3 years.

The causes of inaccuracy of Kaplan–Meier interpretation

The tail of the curve

Towards the end of follow up, the sample sizes are smaller. The occurrence of a given event will have a proportionately higher impact on the survival rate.

Studies omitting the tail of a survival curve may exclude important information. This is particularly important whenever the event more often occurs after longer follow-up times. For example, studies concerning failures as an event after hip arthroplasties should include the tail of the curve along with confidence intervals for the cumulative probabilities, as these failures (i.e. events) usually continue to occur after the end of the study. Presenting the cumulative probabilities with confidence intervals better shows the reliability of the tail (Lettin et al, 1991).

Competing events

A competing event is an event which prevents occurrence of the event of interest. The Kaplan–Meier method assumes that there would be an equal chance for occurrence of the event for all censored subjects after the end of the study. A competing event may cause this assumption to be incorrect if there is no chance of occurrence of the event (i.e. a participating patient's death). The Kaplan–Meier method is therefore not suitable for analyses in the presence of competing events.

Researchers should correct for competing events affecting the data to avoid overestimation of the cumulative risk of the event of interest. In such cases, an alternative statistical method is used such as the sub-distribution hazard model (Gooley et al, 1999).

Future predictions

Kaplan–Meier studies have limited use in predicting future survival rates after the end of the study. Because of this, survival analysis should not be used to extrapolate survival rates beyond the end of a study.

For example, a study by Ritter and Campbell (1987) assessed the survival rates of Charnley hip replacements. A 10-year cumulative survival rate of 92% forecasted a 15-year survival of 88%. The actual 15-year survival rates were found to be 65%.

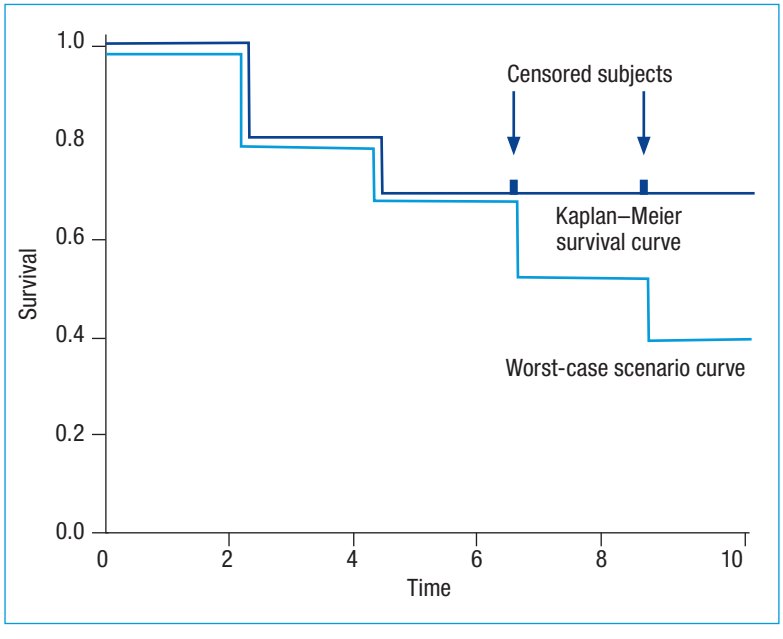

Figure 3. A Kaplan–Meier survival analysis for the cohort in the example is shown with a worst-case scenario curve below. The true Kaplan–Meier estimate would lie somewhere between these two.

KEY POINTS

- When analysing survival rates, ideally a complete actuarial life table and a Kaplan–Meier survival curve should be presented.
- The X-axis shows the serial times while the Y-axis shows the conditional survival probability.
- The cumulative probability is calculated as the product of successive conditional probabilities up to that time point.
- A worst-case survival curve should be plotted assuming that all cases lost to follow up have had the event of interest.
- Confidence limits should be plotted, especially at the tail of the curve.
- The Kaplan–Meier method is not suitable for use in the presence of competing events.

A further example of a Kaplan–Meier curve

The study

Figure 3 shows the Kaplan–Meier curve of a study describing survivorship of a new total hip replacement prosthesis. The event of interest is failure of the prothesis. The study recruited 10 patients and aimed for a 10-year follow-up period.

The patients had their operations at different dates (secular dates). For the purposes of the study, the follow up commenced at point zero (serial time) and was measured in years.

The survival probability

Survival probability includes calculation of the conditional probability and the cumulative probability and are discussed below.

The conditional probability

This is calculated after each event. For example, if after 2 years, two subjects out of the 10 had failure, then the conditional probability would be 0.8 (i.e. failure probability is 2/10) surviving and the number 'at risk' would have decreased to 8. After 5 years from the start of the study, another two participants had failure, then the conditional probability would be 0.75 (i.e. failure probability is 2/8), and the number at risk would have declined to 6 (i.e. six participants remaining).

The cumulative probability

This is calculated as the product of successive conditional probabilities up to that time point. For example, after 5 years from the start of the study, the cumulative probability would be 0.6 which is the product of the previous conditional probabilities (i.e. 0.8 x 0.75).

The incidence rate

From the above example, we can see that the incidence rate for failures after 5 years would be 0.4 (4/10). However, the cumulative probability or probability of survival until 5-year follow up in a decaying population would be different. The cumulative probability for survival at this point was 0.6.

The tail of the study

The remaining number at risk would be low as a result of the successive occurrence of failures and might not be reliable. *Figure 3* shows the addition of a worst-case scenario curve. The real cumulative probability lies between these two curves. In this example, the cumulative probability for the worst-case curve is 0.4 while that for the primary curve is 0.6. The real cumulative probability should lie between these two values.

Conclusions

Actuarial life tables and Kaplan–Meier curves are widely used methods of survival analysis and allow useful values to be calculated for a given data set. Understanding these methods allows better comprehension of the medical literature. **BJHM**

The authors would like to thank Mr Chris Jones for his help in the preparation of the manuscript.
Conflict of interest: none.

Akl EA, Briel M, You JJ, Lamontagne F et al. LOST to follow-up Information in Trials (LOST-IT): a protocol on the potential impact. Trials. 2009 Dec;10(1):40. https://doi.org/10.1186/1745-6215-10-40

Drüeke TB, Locatelli F, Clyne N et al; CREATE Investigators. Normalization of hemoglobin level in patients with chronic kidney disease and anemia. N Engl J Med. 2006 Nov 16;355(20):2071–2084. https://doi.org/10.1056/NEJMoa062276

Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Stat Med. 1999 Mar 30;18(6):695–706. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer Statistics, 2008. CA Cancer J Clin. 2008 Jan 28;58(2):71–96. https://doi.org/10.3322/CA.2007.0010

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958 Jun;53(282):457–481. https://doi.org/10.1080/01621459.1958.10501452

Lettin AW, Ware HS, Morris RW. Survivorship analysis and confidence intervals. An assessment with reference to the Stanmore total knee replacement. J Bone Joint Surg Br. 1991 Sep;73-B(5):729–731. https://doi.org/10.1302/0301-620X.73B5.1894659

Marubini E, Valsecchi MG. 1958. Analyzing Survival Data from Clinical Trials and Observational Studies. 18th edn. Chichester (UK), John Wiley and Sons:41–48.

Murray DW, Carr AJ, Bulstrode C. Survival analysis of joint replacements. J Bone Joint Surg Br. 1993 Sep;75-B(5):697–704. https://doi.org/10.1302/0301-620X.75B5.8376423

Peto R, Pike MC, Armitage P et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer. 1977 Jan;35(1):1–39. https://doi.org/10.1038/bjc.1977.1

Rich JT, Neely JG, Paniello RC, Voelker CCJ, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngol Head Neck Surg. 2010 Sep;143(3):331–336. https://doi.org/10.1016/j.otohns.2010.05.007

Ritter MA, Campbell ED. Long-term comparison of the Charnley, Muller, and trapezoidal-28 total hip prostheses. J Arthroplasty. 1987 Jan;2(4):299–308. https://doi.org/10.1016/S0883-5403(87)80063-3

Stel VS, van Dijk PCW, van Manen JG et al. Prevalence of co-morbidity in different European RRT populations and its effect on access to renal transplantation. Nephrol Dial Transplant. 2005 Dec 01;20(12):2803–2811. https://doi.org/10.1093/ndt/gfi099

Vervölgyi E, Kromp M, Skipka G, Bender R, Kaiser T. Reporting of loss to follow-up information in randomised controlled trials with time-to-event outcomes: a literature survey. BMC Med Res Methodol. 2011 Dec;11(1):130. https://doi.org/10.1186/1471-2288-11-130

Wanner C, Krane V, März W, Olschewski M, Mann JFE, Ruf G, Ritz E; German Diabetes and Dialysis Study Investigators. Atorvastatin in patients with type 2 diabetes mellitus undergoing hemodialysis. N Engl J Med. 2005 Jul 21;353(3):238–248. https://doi.org/10.1056/NEJMoa043545