EXAMPLE (continued)

Group 2 (placebo)

$t_{(j)}$	n_j	m_j	q_j	$\hat{S}(t_{(j)})$
0	21	0	0	1
1	21	2	0	19/21 = .90
2	19	2	0	17/21 = .81
3	17	1	0	16/21 = .76
4	16	2	0	14/21 = .67
5	14	2	0	12/21 = .57
8	12	4	0	8/21 = .38
11	8	2	0	6/21 = .29
12	6	2	0	4/21 = .19
15	4	1	0	3/21 = .14
17	3	1	0	2/21 = .10
22	2	1	0	1/21 = .05
23	1	1	0	0/21 = .00

Thus, considering the group 2 data, the probability of surviving past zero is unity, as it will always be for any data set.

Next, the probability of surviving past the first ordered failure time of one week is given by 19/21 or (.91) because 2 people failed at one week, so that 19 people from the original 21 remain as survivors past one week.

Similarly, the next probability concerns subjects surviving past two weeks, which is 17/21 (or .81) because 2 subjects failed at one week and 2 subjects failed at two weeks leaving 17 out of the original 21 subjects surviving past two weeks.

$$\hat{S}(t_{(j)}) = \frac{\#\text{ surviving past } t_{(j)}}{21}$$

The remaining survival probabilities in the table are computed in the same manner; that is, we count the number of subjects surviving past the specified time being considered and divide this number by 21, the number of subjects at the start of follow-up.

No censorship in group 2

Alternative formula: KM approach

Recall that no subject in group 2 was censored, so the q column for group 2 consists entirely of zeros. If some of the q's had been nonzero, an alternative formula for computing survival probabilities would be needed. This alternative formula is called the Kaplan–Meier (KM) approach and can be illustrated using the group 2 data even though all values of q are zero.

EXAMPLE

$$\hat{S}(4) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16} = \frac{14}{21} = .67$$

$$\Pr(T > t_{(j)} \mid T \geq t_{(j)})$$

For example, an alternative way to calculate the survival probability of exceeding four weeks for the group 2 data can be written using the KM formula shown here. This formula involves the product of terms each of which is a conditional probability of the type shown here. That is, each term in the product is the probability of exceeding a specific ordered failure time $t_{(j)}$ given that a subject survives up to that failure time.

EXAMPLE (continued)

$$\hat{S}(4) = 1 \times \left(\frac{19}{21}\right) \times \frac{17}{19} \times \left(\frac{16}{17}\right) \times \frac{14}{16} = \frac{14}{21} = .67$$

$$\frac{19}{21} = \Pr(T > 1 \mid T \geq 1)$$

$$\frac{16}{17} = \Pr(T > 3 \mid T \geq 3)$$

17 = # in risk set at week 3

Thus, in the KM formula for survival past four weeks, the term 19/21 gives the probability of surviving past the first ordered failure time, one week, given survival up to the first week. Note that all 21 persons in group 2 survived up to one week, but that 2 failed at one week, leaving 19 persons surviving past one week.

Similarly, the term 16/17 gives the probability of surviving past the third ordered failure time at week 3, given survival up to week 3. There were 17 persons who survived up to week 3 and one of these then failed, leaving 16 survivors past week 3. Note that the 17 persons in the denominator represents the number in the risk set at week 3.

$$\hat{S}(4) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \left(\frac{14}{16}\right)$$

$$\hat{S}(8) = 1 \times \frac{19}{21} \times \frac{17}{19} \times \frac{16}{17} \times \frac{14}{16} \times \frac{12}{14} \times \left(\frac{8}{12}\right)$$

KM formula = product limit formula

Notice that the product terms in the KM formula for surviving past four weeks stop at the fourth week with the component 14/16. Similarly, the KM formula for surviving past eight weeks stops at the eighth week.

More generally, any KM formula for a survival probability is limited to product terms up to the survival week being specified. That is why the KM formula is often referred to as a "product-limit" formula.

Next: Group 1

Next, we consider the KM formula for the data from group 1, where there are several censored observations.

Group 1 (treatment)

$t_{(f)}$	n_f	m_f	q_f	$\hat{S}(t_{(f)})$
0	21	0	0	1
6	21	3	1	$1 \times \left(\frac{18}{21}\right) = .8571$
7	17	1	1	$.8571 \times \left(\frac{16}{17}\right) = .8067$
10	15	1	2	$.8067 \times \frac{14}{15} = .7529$
13	12	1	0	$.7529 \times \frac{11}{12} = .6902$
16	11	1	3	$.6902 \times \frac{10}{11} = .6275$
22	7	1	0	$.6275 \times \frac{6}{7} = .5378$
23	6	1	5	$.5378 \times \frac{5}{6} = .4482$

The estimated survival probabilities obtained using the KM formula are shown here for group 1.

The first survival estimate on the list is $\hat{S}(0) = 1$, as it will always be, because this gives the probability of surviving past time zero.

The other survival estimates are calculated by multiplying the estimate for the immediately preceding failure time by a fraction. For example, the fraction is 18/21 for surviving past week 6, because 21 subjects remain up to week 6 and 3 of these subjects fail to survive past week 6. The fraction is 16/17 for surviving past week 7, because 17 people remain up to week 7 and one of these fails to survive past week 7. The other fractions are calculated similarly.

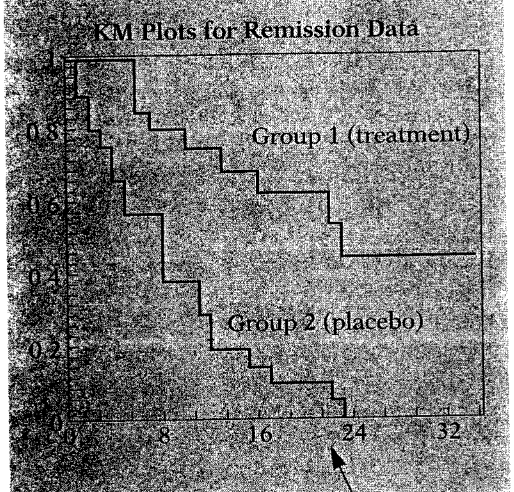
EXAMPLE (continued)

Fraction at $t_{(j)}$: $\Pr(T > t_{(j)} \mid T \geq t_{(j)})$

Not available at $t_{(j)}$: failed prior to $t_{(j)}$
or
censored prior to $t_{(j)}$
↖ group 1 only

For a specified failure time $t_{(j)}$, the fraction may be generally expressed as the conditional probability of surviving past time $t_{(j)}$, given availability (i.e., in the risk set) at time $t_{(j)}$. This is exactly the same formula that we previously used to calculate each product term in the product limit formula used for the group 2 data.

Note that a subject might not be available at time $t_{(j)}$ for one of two reasons: (1) either the subject has failed prior to $t_{(j)}$, or (2) the subject has been censored prior to $t_{(j)}$. Group 1 has censored observations, whereas group 2 does not. Thus, for group 1, censored observations have to be taken into account when determining the number available at $t_{(j)}$.



Plots of the KM curves for groups 1 and 2 are shown here on the same graph. Notice that the KM curve for group 1 is consistently higher than the KM curve for group 2. These figures indicate that group 1, which is the treatment group, has better survival prognosis than group 2, the placebo group. Moreover, as the number of weeks increases, the two curves appear to get farther apart, suggesting that the beneficial effects of the treatment over the placebo are greater the longer one stays in remission.

Obtain KM plots from computer package, e.g., SAS, BMD, EGRET, SPIDA

The KM plots shown above can be easily obtained from most computer packages that perform survival analysis, including SAS, BMD, and EGRET. All the user needs to do is provide a KM computer program with the basic data layout and then provide appropriate commands to obtain plots. The above plots were obtained using the SPIDA package (from Macquarie University in Sydney, Australia). This package will also be used later to present results from log–rank tests.

III. General Features of KM Curves

General KM formula:

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \times \hat{\Pr}(T > t_{(j)} \mid T \geq t_{(j)})$$

The general formula for a KM survival probability at failure time $t_{(j)}$ is shown here. This formula gives the probability of surviving past the previous failure time $t_{(j-1)}$, multiplied by the conditional probability of surviving past time $t_{(j)}$, given survival to *at least* time $t_{(j)}$.