0.502. In calculating the hazard ratio, it is always the case that HR = 1.0 means that there is no difference in the underlying hazard rates of the two groups.

The *P* value corresponding to the significance of separation of the curves from arm A and arm B, from this particular Kaplan–Meier plot was $P = 0.001$. The *P* value is applied for interpreting the experiment as follows. It means that the probability of observing a result as extreme as, or more extreme than, the one actually observed from chance alone is one in one thousand. If the *P* value had been 0.01, it would have meant that the probability of observing a result as extreme as, or more extreme than, the one actually observed from chance alone is one in 100. Because the value of 0.001 is less than 0.05, this means that we can reject the null hypothesis. Holm et al. (16) expressly stated that, "$P < 0.05$ was considered statistically significant." By convention, in the context of handling *P* values, this number (0.05) is called the *alpha value*.

c. Censoring Data

Data on any given subject are "censored" when a subject drops out of the clinical trial. Data on any subjects still alive when a clinical trial on cancer has come to its end, and when the clinical trial has been formally concluded, are also censored. For subjects still alive at the end of an oncology clinical trial, the event that is usually of interest (death) has not occurred, and for this reason the subject is censored.

The term censored means that the exact date of the subject's death is not marked by a downward step on the Kaplan–Meier plot, and is not used for calculating the fraction of surviving study subjects. Where a study subject is censored, this may be indicated on the Kaplan–Meier plot by way of a tick mark or dot. Bland and Altman (17) described the Kaplan–Meier plot as, "the 'curve' is a step function, with sudden changes in the estimated probability corresponding to times at which an event was observed. The times of the censored data are indicated by short vertical lines." Where a subject is censored, a tick mark or dot is shown on the graph.

Generally, subjects are censored when they are lost to the study, but it is not advisable to censor subjects for problems that are less severe, for example, failure to adhere to the drug schedule. If a subject is too ill to travel to the clinic for an infusion of an anticancer drug, and if that subject is censored, the act of censoring that particular subject may introduce bias into the calculations and analysis. Please consider the following hypothetical. In this hypothetical, an experimental drug is ineffective against breast cancer, except for a minority of people in the general population with a rare mutation in the epidermal growth factor gene. Now, please imagine that the health of most of the study subjects deteriorates to the point where they can no longer come to the clinic, and where the investigator decides to censor data from the subjects. In this hypothetical, only a fraction of the subjects—perhaps 5% of the total subjects enrolled in the study—having the rare mutation will feel good enough to come to the clinical for more treatment. In this hypothetical, the subset of study subjects with the rare mutation generally feel good enough to travel to the clinic for scheduled tests and drug doses. The result of the censoring will be as follows. The result will be that the drug is discovered to be dramatically effective against breast cancer. But this will be a misleading and artifactual result because, in fact, the drug is only effective in 5% of the subjects (the subjects with the mutation).

[16]Holm C, Rayala S, Jirström K, Stål O, Kumar R, Landberg G. Association between Pak1 expression and subcellular localization and tamoxifen resistance in breast cancer patients. J. Natl. Cancer Inst. 2006;98:671–80.

[17]Bland JM, Altman DG. Survival probabilities (the Kaplan–Meier method). Brit. Med. J. 1998;317:1572.