Author's response to reviews

Title: Minimum follow-up time required for the estimation of statistical cure of cancer patients: verification using data from 42 cancer sites in the SEER database

Authors:

Patricia Tai (ptai@scf.sk.ca)
Edward Yu (Edward.Yu@lrc.on.ca)
Gabor Cserni (cserni@freemail.hu)
Georges Vlastos (Georges.Vlastos@hcuge.ch)
Melanie Royce (MRoyce@salud.unm.edu)
Ian Kunkler (iankunkler@yahoo.com)
Vincent Vinh-Hung (conrvhgv@az.vub.ac.be)

Version: 2 Date: 6 December 2004

Author's response to reviews:

To the Editor:

Respond to Reviewer 1 (Angela Mariotto):
Major Points:
1. The two registries (Detroit and Connecticut) were used just for illustration of the method. In the discussion it was mentioned that the results were consistent with the nine registries.
2. Patients of all ages were included except for those cancer sites with high frequency of elderly patients. The survival times of patients who died of specific cancer are tested for lognormality. If patients of age greater than 60 were excluded, the lognormal distribution of survival time would be biased.
3. The 1973-1992 data were used so that the data are not out-dated. This is a concern of one of the Reviewers. For a lognormal distribution of the survival time of those patients died of the specific cancer, there is only a very small proportion dying at the tail of follow-up (see Figure 1). So it would not cause much change to the distribution, even with only 7 years of follow-up to 1999 at the tail. For those cancer sites with threshold years longer than 24 years, the 1973-1977 data were used.
4. The age group and year of diagnosis are added to the two Tables.
5. Definition and interpretation of the threshold year are added to the text.
6. The SEER data of survival rates are given as relative survival rates. The survival rates given in this paper are cause-specific survival rates. So they are different.
7. M and S in Table 1 are transformed to median and multiplicative standard deviation of the survival time.

Respond to Reviewer 2 (Claudia Spix):
Major Points:
1. The maximum likelihood estimation gives the mean and standard deviation of the log(survival time) and the p-value. However the minimum chi-square test gives the mean and standard deviation of the log(survival time) at the minimum chi-square value and at the maximum p-value. The test statistic was minimized by varying the parameters and the p-value gave the significance of the test. So the values given are better fit to the data. The class intervals were in the powers of 2 in months of the survival time, such as 0-2, >2-4, >4-8, >8-16, and so on. The minimum chi-square test is used to test if there is no difference between the observed data distribution and the lognormal distribution. It is rejected if P<0.05. All the p-values in Table 1 are above 0.10.
2. The tests were carried out for a large number of patients and for a small number of patients. The results were good for both large and small numbers.
3. Generally cancer-specific death rates underestimate the mortality associated with a diagnosis of the specific cancer, because some patients died of other causes. (Ref. Brown BW, Brauner C, Minnnotte MC. Noncancer deaths in white adult cancer patients. J Natl Cancer Inst. 1993;85:979-987.) SEER is a set of geographically defined, population-based, central cancer registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Registry data are submitted electronically to the NCI on a biannual basis, and the NCI makes the data available for analysis. The SEER Program is considered the standard for quality among cancer registries around the world. Quality control has been an integral part of SEER since its inception. Every year, studies are conducted in SEER areas to evaluate the quality and completeness of the data being reported.
4. The results were justified by comparing with the estimation using Kaplan-Meier method. At the threshold
year, the statistical cure rates estimated for 40 cancer sites were found to match the actuarial long-term survival rates estimated by the Kaplan-Meier method within six percentage points.

5. The present study only estimates the time to wait before statistical cure of the cancer patients from the specific cancer. It does not estimate the cure rates. There are other methods to estimate cure rates, which is a concern of one of the Reviewers.

6. Same as Point 3, generally cancer-specific death rates underestimate the mortality associated with a diagnosis of the specific cancer.

Minor Points:
1. Gamel and Vogel (Ref. Gamel JW, Vogel RL. Comparison of parametric and non-parametric survival methods using simulated clinical data. Stat Med. 1997;16:1629-1643.) have compared the advantage of lognormal distribution over other distributions such as Weibull and log logit.
2. Proportion of survivors and proportion of patients cured are fractions, while survival rate and cure rate are in percentages.
3. We applied the results derived for the lognormal distribution presented by Limpert et al.[24].
4. The estimated threshold years are specific for certain cohorts under study. For new cohorts of recent years, new estimations are needed.
5. The paragraphs of comparing the consistence with other references are re-written for the cancer sites.

Response to Reviewer 3 (Colin Mathers):

Major Points:
1. The definition of the threshold year, $t$, is added to the introduction. The choice of 2.25% follows directly by applying the result derived by Limpert et al.[24], when $t$ is equal to $m^* x (s^*)^2$. The follow-up time for the cumulative predicted risk of dying from the specific cancer becoming less than 2.25% is close enough to cure from that specific cancer.
2. The risk of cancer death and the risk of other cause of death are very fluctuating. It is very difficult to find a threshold where they are stably equal for a period of time.
3. Graphical method to show goodness of fit is subjective and not quantitative. However the minimum chi-square test gives the mean and standard deviation of the log(survival time) at the minimum chi-square value and at the maximum p-value. Gamel and Vogel (Ref. Gamel JW, Vogel RL. Comparison of parametric and non-parametric survival methods using simulated clinical data. Stat Med. 1997;16:1629-1643.) have compared the advantage of lognormal distribution over other distributions such as Weibull and log logit.
4. If there are more patients dying due to other causes than dying of the specific cancer, then the cause-specific lognormal distribution will be biased. Hence it is not expected that all cancer-specific survival time distribution will follow a lognormal distribution.
5. The results were justified by comparing with the estimation using Kaplan-Meier method. At the threshold year, the statistical cure rates estimated for 40 cancer sites were found to match the actuarial long-term survival rates estimated by the Kaplan-Meier method within six percentage points. The present study only estimates the time to wait before statistical cure of the cancer patients from the specific cancer. It does not predict the cure rates. The other reviewer suggests that there are other ways to predict the survival rates. Sc simulations are beyond the scope of the present study.