Reviewer's report

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

Version: 1 Date: 16 November 2004

Reviewer: Angela Mariotto

Reviewer's report:

General

This paper presents a method to estimate the minimum follow-up time required (threshold parameter) for the estimation of statistical cure. The method is based on fitting a log-normal distribution to the times of cancer-specific death. Only patients who died of the specific cancer within the window of observed time are considered in the calculation of the threshold time. Because censored times are not taken into account it is possible that the threshold time estimate might be biased and underestimated in a number of situations:

1) The longest survival or time to death observed is 26 years. All the results are conditioned on a maximum follow-up time of 26-years.
2) For most of the cancer sites the analysis were based on cases diagnosed between 1973 and 1992 with follow-up up to 1999. For cases diagnosed in 1992 only patients who die in 1999 or before contributed to the threshold parameter estimation. Cases that would die after 1999 and would contribute to a longer tail of the log-normal distribution did not enter the analysis. This might bias estimates.
3) Only for some cancer sites age was restricted to patients 60 and younger. For most cancer sites it is not clear which ages were considered. The consideration of older patients can also bias estimates of the threshold time. Older patients die more frequently of other causes than younger patients and thus more patients dying within a shorter time from diagnosis will enter analysis.

Standard survival analysis takes into account censoring either due to end of follow-up or death of other causes. For cancer sites with very short survival such as pancreas the method proposed is accurate because most of the deaths are observed within the duration of the study. However for other cancer sites such as those in which the threshold is close to 20 years this is not true and there might be some biases in the estimates of the threshold time. In order to estimate the bias it is important to do some sensitivity analysis. Estimates should be compared by estimating the threshold parameters restricting cases to those diagnosed during 1973-1977 and restricting age to younger patients.

-------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. It is not clear why only data from Detroit and Connecticut registries were used and not the other SEER registries. The SEER 9 registries are Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah. Data are available for cases diagnosed from 1973 and later for these registries with the exception of Seattle-Puget Sound and Atlanta. The Seattle-Puget Sound and Atlanta registries joined the SEER program in 1974 and 1975, respectively.
2. Analysis should be restricted to patients 60 or younger because of bias due to deaths of other
causes. Sensitivity analysis could be carried out to investigate the effect of deaths due to other causes.
3. Sensitivity or restricting data to 1973-1977 should be performed in all cases that used 1973-1992 data.
4. Tables should clearly state which age groups and year of diagnosis were considered for each cancer site.
5. Definition and interpretation of threshold should be given in the text.

6. The 20-year survival rate for cervix uteri in the SEER data is 61% (see reference 27 from their paper) The paper gives a 92.7% which seems to be very high.
7. The values of M and S in Table 1 should be transformed to mean survival and standard deviation of the survival time, to have more interpretable values.

---

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests.