Author's response to reviews

Title: Different Regression Equations Relate Age to the Incidence of Lauren Types 1 and 2 Stomach Cancer in the SEER database: These Equations Are Unaffected by Sex or Race

Authors:

Mitchell S Wachtel MD (mitchell.wachtel@ttuhsc.edu)
Yan Zhang MD (yan.zhang@ttuhsc.edu)
Maurizio Chiriva-Internati PhD (maurizio.chiriva@ttuhsc.edu)
Eldo E Frezza MD, MBA (eldo.frezza@ttuhsc.edu)

Version: 3 Date: 2 February 2006

Author's response to reviews: see over
Reviewer’s report
Different regression equations relate age to the incidence of stomach cancer in the SEER database for Lauren type 1 and Lauren Type 2 stomach cancers and for Asians and non-Asians; these age relationships are the same for men and women.

Version: 2 Date: 24 January 2006
Reviewer: Patricia Tai

Reviewer’s report: General
This journal is intended for oncologists, so explain more about the statistical methods involved.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
1) Under Model adequacy, explain what is the meaning and use of dffit and Cook’s distance.

The sentence now reads:
A data point was considered an outlier if its studentized residual, calculated by R, was greater than 3; dffits (a measure which gives greater weight to outlying observations) and Cook’s distance (a measure of impact of the respective case on the regression equation), calculated by R were used to assess leverage; outliers that are found to have high leverage, large dffits and/or Cook’s distances, are considered bad leverage points and are removed from the analysis.

2) In Table 1, what is the meaning of Sex (no total)?
The words ‘(no total)’ have been removed.

3) In Table 3, R2 should be R square with 2 superscripted.
The superscript has been added.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
1) In the Title, change “Lauren Type” to “Lauren Types”.
The suggestion has been adopted.

2) Under Data transformation, “Ppreliminary” should be “Preliminary”.
The suggestion has been adopted.

3) Under Final model, put “residual sum of squares” before RSS, and put “main effects” before ME.
The sentence now reads:
The final model, displayed in Table 8, included the four other main effects (ME) covariates and the four interaction covariates found to significantly reduce residual sum of squares (RSS).

4) Under Final model, in Lauren type 2 equation, should be log(ca).
The suggestion has been adopted.

5) For many reference quotes, put comma or full-stop after the reference number, e.g. Levi [13], and Koirea [31].
Bracketed references were considered in the way that superscripted references once were considered. Hence, if the immediately preceding word would have required punctuation, the punctuation mark was placed immediately after the right bracket. This wise suggestion required many changes. Spaces between the right brackets and punctuation marks were removed. Some punctuation marks were added. Some punctuation marks were removed. If the editor’s policy is to place punctuation marks after each bracketed reference, this will assuredly be done, but it might confuse the reader.

6) Many Figure legends are missing.
Figure legends now read:
Figure 1a. Plot of the logarithms of cancer rates, denoted as log(ca), as a function of age in years for Asian men. Red references Lauren type 1 gastric cancer. Blue references Lauren type 2 gastric cancer. Lines represent predicted values.

Figure 1b. Plot of the logarithms of cancer rates, denoted as log(ca), as a function of age in years for Asian women. Red references Lauren type 1 gastric cancer. Blue references Lauren type 2 gastric cancer. Lines represent predicted values.

Figure 1c. Plot of the logarithms of cancer rates, denoted as log(ca), as a function of age in years for non-Asian men. Red references Lauren type 1 gastric cancer. Blue references Lauren type 2 gastric cancer. Lines represent predicted values.

Figure 1d. Plot of the logarithms of cancer rates, denoted as log(ca), as a function of age in years for non-Asian women. Red references Lauren type 1 gastric cancer. Blue references Lauren type 2 gastric cancer. Lines represent predicted values.

7) The y-axis label in the figures should be log(ca), the " ) " is missing. The suggestion has been adopted.

8) In Methods, second last sentence of first paragraph: “The two five-year....” should insert the hyphen. The suggestion has been adopted.

Reviewer’s report

Different regression equations relate age to the incidence of stomach cancer in the SEER database for Lauren type 1 and Lauren Type 2 stomach cancers and for asians and non-asians; these age relationships are the same for men and women.

Version: 2 Date: 25 January 2006

Reviewer: P W James

Reviewer’s report:
- Major Compulsory Revisions
  I have none. I think the statistical methodology is much better now and I suggest just a few minor revisions.
- Minor Essential Revisions
  In 'Model selection' section the list of five covariates should be referenced as a), b), c), d), e).

The suggestion has been adopted.

I notice that under Final Model you have included 'age squared' in the group of variables which could interact with Lauren Type but you have not mentioned 'age squared' in the Summary of model building. Please make clear in this section if you considered non-linear functions of age such as 'age squared' in the model.

The words 'age squared' were eliminated; Professor James’ wise suggestion as to logarithms eliminated the need to consider the non-linear functions and higher order interactions of explanatory variables. In R, the term ‘log(x)’ references the natural logarithm, not the base 10 logarithm; the text was changed throughout to reflect the fact that natural logarithms were taken. To the model selection section was added the sentence:

Neither higher-order explanatory variable interactions nor non-linear relationships of explanatory variables to the response variable were considered in this analysis.

The Lauren type 2 equation should start as 'log(ca)= ' for consistency with Lauren Type 1 equation.

The suggestion has been adopted.

Figures 1, 2 and 3 need a legend to say that red means type 1 and blue means type 2. The authors may need to consult the editor about the preferred format for tables.

The suggestion has been adopted.

I suggest that the Tables need a few more horizontal lines; for example Table 4 needs a line at the top and the bottom.

The suggestion has been adopted.