Author's response to reviews

Title: Tumor location and patient characteristics of colon and rectal adenocarcinomas associating survival with TNM classes

Authors:

  Kari Hemminki (k.hemminki@dkfz.de)
  Irene Santi (i.santi@dkfz.de)
  Marianne Weires (m.weires@dkfz.de)
  Hauke Thomsen (h.thomsen@dkfz.de)
  Jan Sundquist (jan.sundquist@med.lu.se)
  Justo Lorenzo Bermejo (J.Lorenzo@Dkfz-Heidelberg.de)

Version: 2 Date: 24 September 2010

Author's response to reviews:

Dear Professor Clarke,

Thank you for the helpful comments on our paper. We have now modified it by considering all the points raised be the reviewers. The changes are underlined in the revised manuscript. As a major change we have used and updated version of the Swedish Cancer Registry for study (years 2004 to 2009 for the TNM part and years 2000-2008 for the survival part), as these data became available during the review process. The figures are new but no interpretation has changed. Also the language has been attended. We hope that the revision would be acceptable for publication.

Sincerely yours,
Kari Hemminki

DETAILED COMMENTS

Reviewer's report
Title: Tumor location and patient characteristics of colon and rectal adenocarcinomas associating survival and with TNM classes
Version: 1 Date: 16 June 2010
Reviewer: Peter Gibbs

Reviewer's report:
The authors have attempted to address a number of questions utilising data from the Swedish Cancer database, in particular examining the relationship between age, tumour site and survival.

Minor essential revisions
1. I would be interested in more information as to how some of the data is
obtained as I am uncertain as to how accurate the data in the registry is. Is the quality of the data ever audited? In particular I would be interested to know how data regarding M stage and family history were obtained by the database, and with respect to M stage is CT staging at diagnosis routine in Sweden? A family history of colorectal cancer in only 5% of patients seems a little low.

>>> Data on TNM classification was added to middle of p.5. Cancer registry only has data on the basis of TNM: clinical or pathological (text in first 2 paragraphs of Discussion). Family history covers only the years of the survival study, 2000-2008 (end of first paragraph, p. 7).

2. The more advanced stage at diagnosis in young patients is presumed due to more aggressive tumours, but could be due to delayed diagnosis or delayed presentation. This should be discussed.

>>> This point was added: line 8 from the end of p. 9.

3. The limitations of the work have not been addressed. A paragraph or two related to this would be appropriate.

>>> Text was added to the first 2 paragraphs of Discussion.

4. The title is a little confusing, an alternative title should be attempted.

>>> Modified.

Reviewer's report
Title: Tumor location and patient characteristics of colon and rectal adenocarcinomas associating survival and with TNM classes
Version: 1 Date: 17 June 2010
Reviewer: Owen Franklin Dent
Reviewer's report:
Major Compulsory Revisions
1. This study is poorly conceptualized, poorly prepared and poorly written and the findings are not novel.

>>> Dr. Dent makes a major point (no 3 and 20) of ‘normal mortality’. Unfortunately he does not appear to known that in survival analysis ‘normal mortality’ is the reference. As some other readers may not know this we added a sentence, top of p. 6: Note that the HR compares the death rate of CRC patients to that of the population considering the above variables, including age.

>>> The following points 2, 4-7 are related to the presentation, which we have improved.
2. The title of the paper is obscure.

3. Regarding the first four lines of the abstract: (a) CRC tends strongly to be a disease of old people and old people tend to die because of what demographers call "normal mortality" – so the reason for the association between age and poor survival in CRC is not unknown, it is quite obvious, (b) the second sentence is unrelated to both the first and the third, (c) the third sentence should indicate the research question and state the aim of the study but does not, rather it says simply that the study provides data.

4. In the first sentence of the introduction it is tautological to say that mortality has declined because survival has improved.

5. Page 3 introduces several disparate ideas but does not develop a clear theme leading towards a statement of the general research question and the aim of the study. Also, as in the abstract, it ignores the fact that CRC is a disease of old people who experience normal mortality.

6. The paragraph at the top of page 4 says briefly what was done but does not say why. There is no clear, explicit statement of the aim of the study.

7. In the second line on page 4 you say that you examined data collected over four years (presumably 2002 to 2006?) but about two-thirds of the way down page 4 you say that you retrieved cancer registry data from 1958 to 2006. What was the point of retrieving data from 1958 to 2001?

8. The description of the T, N and M categories at the top of page 5 does not precisely how they were coded in the analyses in Table 3. Without knowing the coding it is not entirely clear how the odds ratios should be interpreted.

9. At the end of the methods section it remains unclear precisely which group of patients you studied and from what time period. For example, if TNM staging was used then this implies postoperative staging which implies that all patients had a resection whereas the passage beginning in line 5 of the second paragraph of page 4 implies that patients were included even if their cancer was inoperable. Furthermore, the first line of the results refers to patients diagnosed since 1990; does this include all diagnoses, whether a resection was performed or not, and why 1990 rather than 1958 or 2002?
10. The variables socioeconomic status, medical region and first degree family history of CRC should be defined explicitly.

>>> Done, last line, p. 5, top of p. 6.

11. The method followed in logistic regression modelling should be explained.

>>> Last paragraph of methods.

12. In the second line of the second paragraph on page 5 "sex" would be a more appropriate word than "gender".

>>> Done.

13. The first paragraph of the results describes a survival analysis but this is not mentioned in the methods section and the reason for it is not specified. The survival outcome variable should be specified along with censoring and the method of survival analysis. Information should be given about the follow-up protocol.

>>> Added, bottom p. 5.

14. Table 1 should be cited in the first paragraph of the results but is not. The heading to Table 1 refers to cause-specific survival but the method by which CRC-specific survival was analysed is not described. If the method was Kaplan-Meier with deaths due to causes other than CRC censored then this would be incorrect. Because a considerable proportion of deaths among patients with CRC are due to causes other than CRC the survival analysis should take account of competing risks. See Putter H. et al. Tutorial in biostatistics: competing risks and multi-state models. Statistics in Medicine 2007;26:2389-2430 and reference 17 in that paper.

>>> Cox cause-specific survival, p. 5.

15. In Tables 1 and 2 "ascendens" should be "ascending" and "decendens" should be "descending" and "sigmodeum" should be "sigmoid".

16. It is not clear what the point of the analysis in Table 1 is. Why focus specifically on left versus right side as the outcome of interest (note that site is not adjusted in respect of the other variables) whereas earlier in the paper you emphasized stage as the variable which might explain the association between older age and poorer survival? Why was stage not adjusted? If it is because you did not have stage data for these patients this should be stated explicitly in the
methods section. If stage was not available for all of these patients then why didn’t you restrict the analysis to the period for which it was available? The whole question of what this analysis is about and why it was done has not been explained.

17. How do the 17,487 patients in Table 2 relate to the 51,172 patients in paragraph 1 of the results? Why did you not do all analyses on the one group of patients?

>>> Numbers have been revised because of updating.

18. At the foot of page 5 you say that Table 2 shows the key variables but surely stage is a key variable, yet it is not included whereas family history, which is peripheral to the very large majority of sporadic CRC, is included.

>>> Revised.

19. The footnote to Table 3 does not say that side was adjusted in the models. Is this an oversight, or if it was not adjusted, why not when all other variables were? You say that sex and family history had no effect in the models, so why have you adjusted for them? Why not reduce the models to adjust for only those variables which had effects, which is the conventional approach?

>>> All the variables shown in Table 3 were considered in the analysis. Sex and family history has had effect in some earlier studies and they are thus of interest.

20. The discussion is confused and confusing. It tries to grapple with too many issues at once without developing a clear progression of ideas. Perhaps the most telling sentence is where you say "The finding of more aggressive tumors among young patients was opposite to the survival data which were better in young patients." yet nothing you have done works towards explaining why this is so. As commented earlier, you have entirely overlooked the issue of normal mortality. Furthermore the sentence "Thus the overall conclusion is that poorer survival of old patients in colon cancer is unlikely to be related to delayed diagnosis." is unjustifiable because none of your analyses concerned diagnostic delay.

However the very last sentence of the paper is entirely correct and this is what you should have been taking into account throughout the study. If your data source does not allow you to do this then it is questionable whether any of your analyses are worthwhile.
Reviewer's report
Title: Tumor location and patient characteristics of colon and rectal adenocarcinomas associating survival and with TNM classes
Version: 1 Date: 29 June 2010
Reviewer: Kho Patricia
Reviewer's report:
This paper presents interesting retrospective data that is hypothesis generating. The major strengths of the paper are the large numbers of patients in the cohorts and well collected data.

Major revision
1. Limitations of the study need to be stated – retrospective analysis, other possible covariants that could explain effects lacking, treatment differences over different time periods between cohort for survival analysis and cohort with TNM data, etc

>>> First 2 paragraphs of Discussion.

Minor revisions
1. The objective/aim of the paper is currently included in Background of abstract. It would be clearer if it is stated under its own separate subheading. The title could be shorter and less ambiguous.

>>> Both revised.

2. It is unclear as when the cohorts start and end. Does the first cohort used for survival analysis start from 1990 to 2002 or 2006? Again it is assumed that the 2nd cohort with TNM data is from 2002 to 2006 due to mention on pg4 that TNM started in 2002 and there is 4 years of data. It would be clearer if categorically defined under “subjects and methods”.

>>> The updated follow-up periods are clearly stated now.

3. Addition of a conclusion after the discussion would be useful to summarize the essential points of the paper.

>>> Done.

4. TNM should be defined when first used in abstract.

>>> Done

5. There are some grammatical revisions required as some sentences are too
long and lack punctuation.

Language has been attended.

Discretionary Revisions
1. It would be more clinically relevant if age groups divided in broader ranges such as <50, 50-64, 64-79, 80+? It might not however make any difference to the final results but these groupings more relevant to clinicians.

>>> No change has been done.

2. It would be interesting to know if there was any missing data for TNM staging.

>>> Added to middle paragraph, p. 5.