Reviewer's report

Title: Delay factors in the diagnosis and treatment of symptomatic colorectal cancer

Version: 1 Date: 26 July 2012

Reviewer: Rosemary Tate

Reviewer's report:

An interesting paper investigating the factors associated with delay in treatment of colorectal cancer.

This paper needs major revision. The results are important and should be published, but in order to justify the conclusions much more detail is needed about both the measures and the (multivariate) methods. All suggestions are for compulsory or minor essential revisions.

General comments on the statistical analysis

The descriptive statistics (i.e. medians IQR’s) seem fine, although the section in the abstract needs rewriting to specify the measures and to detail the (descriptive) methods used.

It is not clear how the measures based on the medical records were extracted (from codes, free text?) or how counts of symptoms were constructed.

The section explaining the multivariate regression is very unclear and needs further explanation.

Also unclear is exactly what is being tested in the Cox regression model.

What questions were being asked for this part of the analysis and why are the events classified as time-dependent since many of them (e.g MRI, CT scan) are part of the diagnosis process which (presumably) occur at the end of the SDI?

The statement “introduced as time-dependent covariates, which allowed consideration of the time at which these occurred” needs to be clarified and also detailed much more in the results section.

Normally with survival analysis the assumptions are tested before committing to a method. Here Cox regression was chosen before the data was obtained, and no results are given for the tests of proportionality.

More detailed comments

Abstract

1. This needs considerable revision

Methods: The measures need to be described more clearly detailing which is the outcome and which are the covariates
Cox analysis (and measures for it) should all be described in the methods section.

The phrase “median symptom diagnosis interval” does not really make sense unless SDI is first defined in the methods section.

Results

“1rst” should be first

Main paper

Methods

2. This is not a descriptive study as it involves statistical modeling and analysis. It is a cross-sectional study where patient interviews as combined with data from patient notes. More detail is needed on how the primary care measurements were extracted. Was this coded data or free text notes? Were the dates of symptom presentation and diagnosis validated – if so how and if not, can you cite refs that check this? This is important – as explained in ref 53?

3. I note that in the original proposal for this work you planned to compare patient-reported and GP-reported symptoms? Although you cite other authors who have done this, there is no mention of this here? This should be reported.

4. For the Cox model, were the assumptions of proportionality tested and if so what was the result? This needs to be reported. (Alternatively a Poisson model may be more appropriate)

5. Many different types of variables are mentioned, but little detail of how some these were obtained, and/or aggregated, is provided in either this paper or the protocol paper. Some variables e.g. number of visits number of symptoms are only mentioned in the results but are not defined in the Methods. Some have different names in different tables e.g. “Referral mention of CRC” in table 4 is renamed as “CRC suspicion in referral letter” in table 6. (I assume these refer to the same thing, or is this a new variable introduced for the Cox regression?) Much more detail (and consistency) is needed. Some of this detail could be given in an appendix if appropriate.

Results

6. The (bivariate) tables are well presented, although some variables that are shown in Table 6 seem to be missing – e.g. abdominal occlusion, CRC suspicion why is this? Please include, together with the numbers involved.

7. The result that family history of cancer is positively associated with delay seems surprising and counter-intuitive. This result needs checking and further investigation.

I note that 45% of patients reported this family history of cancer, which seems rather high – although again this variable is not described, so I can only assume it must relate to non-immediate family and all types of cancer (including very
common benign skin cancers?) - another measure that needs more detailed explanation.

8. The results of the Cox regression are very unclear and I’m not really sure why these are included as so little is said about the results. This is the first time symptom duration has been mentioned – presumably you mean symptom diagnosis interval?

9. In Table 6, all the p-values should be given (NS is not sufficient – do you mean N/A (not applicable)? If so explain why this is so.

   It is not possible to have a p value of <0.000

Statistical Analysis

10. Please explain what you mean by “In order to avoid outliers”?

11. The sentence “Most of our patients have to wait median of 4 months” doesn’t make sense and is incorrect. By definition, half the patients have a value greater or equal to the median. The correct statement is “half (or 50% of) the patients have to wait at least 4 months”.

Conclusions

Although these conclusions may be correct, I cannot have confidence in them until the measures are better explained and explanation of the analysis is improved.

12. I do not understand this sentence “These findings reinforce the premise that prior experience with cancer in relatives or friends results in longer symptom duration [38]”

   This ref is from 1988 are there no more recent ones? This result seems quite implausible.

References

13. The title of ref 53 is incorrect. It should be

   Determining the date of diagnosis – is it a simple matter? The impact of different approaches to dating diagnosis on estimates of delayed care for ovarian cancer in UK primary care

   While this paper briefly mentions non-recording of symptoms this is not what it is about – please find another ref for this. You may like to cite another of our papers that does show that symptoms often go uncoded and are recorded in the free text


   Please check that the other references are correct.
Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.