Reviewer’s report

Title: Young-Onset Colorectal Cancer in the North East of Scotland: Survival, Clinico-Pathological Features and Genetics

Version: 0 Date: 10 Dec 2019

Reviewer: Matty P. Weijenberg

Reviewer's report:

Overall this is a very interesting paper especially the results regarding the very poor reference to genetic testing. In my opinion this could have a more prominent focus in the paper.

Introduction

1. Why only go into Lynch syndrome in more detail and not also into FAP in the introduction? The intro is a bit misleading stating that FAP is defined by MSI.

Methods

2. How common was it for patients not to have notes or insufficient notes?

3. Since when have the new guidelines for genetic referral been used? After 2015? This is mentioned in the discussion, please mention already in the methods and describe differences.

4. Survival was also collected until 2015 and from the NHS Grampian general and genetic patient records? A bit more explanation is needed about the purpose for these records for a broad audience not necessarily familiar with these records. I assume these are not primarily collected for research purposes. Can you provide any information about loss to follow-up?

5. How is recurrence defined?

Results

6. Could you state a bit more about the characteristics and/or reasons for those excluded from the study (n=73, 17% of population)?

7. In table 1, definition for high risk in the first cell, it is stated 'One affected relative ≤50, one must be first degree relative of another', should the latter not be 'of one another'? Also add 'years' after the ages.
8. It seems that for some patients, the time of follow-up (retrospectively) was shorter than others, since the inclusion period was between 2005 and 2015? For which part of the population did you have at least 5 years of follow-up? There is potentially an overestimation of survival and recurrence. Since different factors were associated with survival, it is interesting to know how age at diagnosis was associated with survival adjusted for Dukes' stage, presentation type (e.g. screening) and sex for example. It is good to not only show the mean survival in table 2, but mean and standard deviation or median and interquartile range. You mention significant differences, but do not show the significance in the table. For sex, it does not corroborate with figure 1 where the Kaplan Meijer curves do not seem to differ significantly between men and women. What do the figures in table 2 look like for those individuals with at least 5 years of follow-up?

9. It would be helpful in table 3 to put the percentages for the totals within age groups to help compare age groups. This is not possible now. For example, in the text, it is mentioned that only 30% of patients aged 50-55 years of age were identified through the screening program, but the percentage is not indicated in the table. Also it is not clear for which comparisons the P-values in the last column are meant. Is this for the comparison between characteristics or also between age groups? It seems to differ in the column. It is also informative to add percentages in the last column. What percentage of the total population had a recurrence for example?

10. It is difficult to assess the recurrence and mortality rate if it is unknown whether there was enough follow-up time for all individuals included (at least 5 yeears) (table 3).

11. The information on being a lifelong vegetarian is a bit odd in the list of characteristics. Was this information collected for all patients? How was this information collected? Maybe this is less relevant for the topic of this paper.

12. With the number of deaths and recurrences, it would be possible and informative to conduct survival analyses adjusted for potential confounders to investigate how the associations with different characteristics were independent of other characteristics. This would also account for time till event.

Discussion

13. Comparing the five-year survival rates in this cohort with the general population is not possible if not all individuals in this cohort were followed-up for at least five years.

14. Although the recent paper on antibiotic use is interesting, it may be a bit over-stretched to use this in the discussion since there is no data on antibiotics use in the current study and the evidence for the relevance of the difference in colon versus rectum cancer incidence is not yet established.

15. "poor statistical significance" is strange wording, preferably use "low power".
16. There is mention of two cohorts in the discussion with regards to the application of the SIGN 2003 guidelines. It is not clear what is meant. The differences between the 2003 and 2011 guidelines should be mentioned earlier in the methods section when the 2003 guidelines are mentioned for the first time.

17. Is the poorer prognosis in those under 40 years compared to those older than 40 years a significant finding in your study? It would be interesting to test this, preferably with correction for other potential prognostic factors.

References

18. Please check references: some are now without an indication of the journal.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:
1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal