Author’s response to reviews

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

Authors:

Sarah Perrott (sarah.perrott.16@abdn.ac.uk)
Kirsten Laurie (kirsten.laurie@nhs.net)
Kirsten Laws (kirstenlaws@nhs.net)
Annie Johnes (ajohnes@nhs.net)
Zosia Miedzybrodzka (zosia@abdn.ac.uk)
Leslie Samuel (leslie.samuel@nhs.net)

Version: 1 Date: 03 Jan 2020

Author’s response to reviews:

Dear Editor,

Thank you for reviewing our manuscript. We greatly appreciate the reviewers’ feedback and comments. We hope you are satisfied with the revisions we have made.

Please see the attached manuscript with the changes highlighted in one of three colours: yellow corresponds to Elizabeth Dennett’s comments, green corresponds to Matty Weijenberg’s comments and comments highlighted in blue are further minor changes we have made independent of the reviewers’ comments.

Please find a point-by-point response to the reviewers’ comments below. We hope the manuscript is now acceptable for publication.

Yours sincerely,

Sarah Perrott

Elizabeth Dennett (Reviewer 1) Comments
Thank you for the opportunity to read this manuscript. We found it to be an interesting review of CRC in younger patients and it goes some way to supporting the argument the age for screening should be reduced in many countries.

However, you need to change the staging through the manuscript.

In his 1948 paper Cuthbert Dukes never described a stage D this was added later, what iteration of the Dukes staging system did you use?. Irrespective of that the Dukes system is no longer used, We am surprised you stuck to it for this manuscript. The TNM should be used and even if you can not accurately T and N and M each tumour you should be able to put the patients into the appropriate Stage 1-4 category not Dukes A- D.

The staging has been changed from the Dukes system to the I-IV numerical staging system. Changes are highlighted in yellow and can be found at:

Methods, line 12, page 4
Methods, line 19, page 4
Results, line 11, page 5
Results, table 2, page 6
Results, line 6, page 7
Results, lines 8-9, page 7
Results, table 3, page 8
Discussion, lines 14-16, page 10
Discussion, lines 9-12, page 11
Lines 47-50 page 5 of the proof

Five year survival rates in those aged less than 40 years was found to be worse (57%) compared to those aged between 40 and 54 years (68% and 62%).

what does the 68% and 62% refer to? the way the sentence reads there should be only one % associated with 'between 40 and 54 years'

It is only when you look at table 2 it makes sense

68% refers to those between 40 and 49 years and 62% for those between 50 and 54 years

This needs to be made clearer in the text or you should put 68-62% not 68% and 62%.
Thank you, 68-62% has replaced this. It is highlighted in yellow (results, line 4, page 6).

Matty P. Weijenberg (Reviewer 2) Comments

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.

Please find an additional sentence about FAP in the background. (Background, lines 10-11, page 3)

Thank you, this sentence now reads as:

'The latter are usually defined by germline mutation in mismatch repair genes in the case of Lynch syndrome or by a germline mutation in the adenomatous polyposis coli (APC) gene for familial adenomatous polyposis (FAP).'

We used ‘usually’ as we are aware there are even rarer hereditary CRC conditions such as Peutz Jeghers syndrome which are out-with the scope of this paper. (Background, lines 4-6, page 3).

Methods

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

As discussed in the results section (see discussion point 6).

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.

A statement has now been added to methods regarding the change in guidelines in 2011 and the differences. With regard to genetic referral for CRC patients, the differences between the 2003 and 2011 guidelines are negligible. Therefore to maintain consistency the 2003 guidelines were used throughout the study. This did not affect the data collection in any way. (Methods, lines 25-26, page 4 and lines 1-2, page 5)
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?

The NHS Grampian general and genetic patient records are kept for the purpose of a patient’s care. Two sentences have been added to clarify this (methods, lines 7-9, page 4). One patient who emigrated was lost to follow-up. A sentence has been added (methods, line 3, page 4).

5. How is recurrence defined?

Disease recurrence was defined by the presence of disease post-treatment on follow-up and imaging (methods, lines 14-15, page 4).

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)?

Please find the reasons for exclusion now listed (results, lines 3-6, page 5). 30 patients had incomplete or insufficient notes.

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.

This has now been reworded and is a lot clearer. ‘Years’ has also been added. (Methods, table 1, page 5)

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?

Data was collected in August 2018, so all patients had at least two years follow up. 67 out of the 345 patients were diagnosed after August 2013, therefore 278 out of 345 (80.6%) patients had at least 5 years follow up. We have added this statement to results:

'All patients had at least two years follow-up and 80.6% of the population had at least five years follow-up.'

(Results, lines 2-3, page 6)
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.

See discussion point 12 regarding Cox Regression analysis.

It is good to not only show the mean survival in table 2, but mean and standard deviation or median and interquartile range.

We appreciate your comment, although standard deviation is not calculated by the SPSS survival analysis. Although standard deviation would be more appropriate, standard error is calculated by SPSS so we have added that to table 2 instead. (Results, table 2, page 6)

You mention significant differences, but do not show the significance in the table.

The p values were shown on figure 1 instead of the table. We have put them in the table too (results, table 2, page 6).

For sex, it does not corroborate with figure 1 where the Kaplan Meijer curves do not seem to differ significantly between men and women.

Thank you, yes the survival between males and females is not significant (p=0.212 – figure 1) so I have therefore removed ‘male sex’ from the sentence: 'In the univariate analysis, the following factors were associated with longer survival: increased age, stage 1 or 2 disease, presentation via screening and male sex.' (results, line 1, page 6)

What do the figures in table 2 look like for those individuals with at least 5 years of follow-up?

80.6% of the population had at least 5 years of follow-up. When the 67 individuals diagnosed after August 2013 (since data was collected in August 2018) who did not have 5 years follow-up are excluded, the values in table 2 remain consistent. For example, overall mean survival is 97.0 (SE±3.7) months and 5YS is 63%.

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.

We agree, ideally it would be helpful to include the percentages of each characteristic within each age group for comparison. To do this we would need to make another separate table to include this data. Although this would be nice, we are not sure it warrants an additional table as one is able calculate the percentage from the data in table 3 if they wish. We could also provide this as supplementary data too if this is of interest.
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.

You are right, they do differ. They were initially meant as a comparison between the characteristics. To avoid confusion we have now removed them from the table. The final column is now ‘total’.

It is also informative to add percentages in the last column. What percentage of the total population had a recurrence for example?

Percentages have been added to the last (total) column. (Results, table 3, pages 7-8).

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 years) (table 3).

Yes, this is true. However, all patients did have at least two years follow-up, which should be sufficient time to assess the recurrence and mortality rate. 80.6% of the population had at least five years follow up.

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.

When a CRC patient is initially seen by oncology in Aberdeen, an extensive history is taken. As red meat consumption is seen as a risk factor for CRC, a patient’s diet is noted on their electronic file. This information should be collected for all patients although there are likely to be inconsistencies. It is irrelevant to the topic however so the statement has now been removed.

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.

We agree, it would be informative to use the Cox Proportional-Hazards model for this. Unfortunately our population size is too small and is likely to yield type I errors.

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.

See discussion point 8. Even excluding the patients who did not have at least five years follow-up, five year survival is still calculated to be the same (63% overall). A sentence has been added to the discussion regarding this. (Discussion, lines 2-4, page 13)
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.

Thank you, the two sentences regarding antibiotic usage have now been removed. These were: 'Interestingly, a recent BMJ case-control study investigating oral antibiotic use and CRC risk found that antibiotic use decreased distal, or rectal, cancer risk, but increased proximal, or colon, cancer risk(17).'

'We found the vast majority of yCRC patients to have a left-sided or rectal tumour, which would question whether increased antibiotic use is a significant risk factor for yCRC(17).'

15. "poor statistical significance" is strange wording, preferably use "low power".

Thank you, this has been changed to 'low power' (discussion, line 2, page 13).

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.

As seen in discussion point 3.

Thank you, we hope we have made this clearer now. Dr Kirsten Laurie gathered data from 2005-2009 in 2015. In 2018, Sarah Perrott gathered the 2010-2015 data, then amalgamated both datasets to have figures over the eleven-year period. The differences between the 2003 and 2011 guidelines are actually negligible with regard to CRC. We have now also mentioned this in the methods section.

(discussion, lines 9-15, page 13)

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 factor.

Unfortunately the population is too small for correction for other potential prognostic factors (see discussion point 12). Only 40 patients were aged less than 40 years and 305 were aged 40-55 years. Mean survival in <40s was 60.9 (SE±6.6) months and 40-55 years was 96.8 (SE±3.6) months. When the log-rank test was performed comparing the four age groups (see table 2), p=0.005 therefore this finding is significant.

References

18. Please check references: some are now without an indication of the journal.
Thank you, these have been fixed now.