

Author's response to reviews

Title: Penetrance of Colorectal Cancer among MLH1/MSH2 Carriers Participating in the Colorectal Cancer Familial Registry in Ontario

Authors:

Yun-Hee Choi (Yun-Hee.Choi@schulich.uwo.ca)
Michelle Cotterchio (Michelle.Cotterchio@cancercare.on.ca)
Gail McKeown-Eyssen (gail.eyssen@utoronto.ca)
Monga Neerav (Neerav.Monga@cancercare.on.ca)
Bharati Bapat (Bapat@lunenfeld.ca)
Kevin Boyd (Kevin.Boyd@cancercare.on.ca)
Steven Gallinger (sgallinger@rogers.com)
John McLaughlin (John.Mclaughlin@cancercare.on.ca)
Melyssa Aronson (maronson@mtsinai.on.ca)
Laurent Briollais (laurent@lunenfeld.ca)

Version: 2 **Date:** 29 July 2009

Author's response to reviews: see over

Wednesday, July 29, 2009

Dear Editor,

We greatly appreciate the reviewers' comments and we revised our manuscript following the suggested advice. A point-by-point reply is attached below.

Reviewer: Yvonne Hendriks

The study design is therefore promising. Unfortunately the resulting cancer risks that are calculated have very large confidence intervals, especially those shown in table 2. Intervals for example from 0 to 70 and 9 to 83 are shown. In table 3 the intervals are somewhat smaller, after another analysis was performed. My knowledge of statistics is not sufficient to determine whether this is a correct analysis. I think that the confidence intervals are too wide to state that the cancer risks calculated in this study are lower than mentioned in earlier studies. The results are definitely too weak to define clinical management in carriers. And this should not be suggested in the manuscript.

The penetrance estimates in Table 2 are based on Kaplan-Meier (KM) analysis, a classical method for analyzing this type of data. We agree that it provides large confidence intervals (CIs) and was given just for comparison with the modified segregation analysis. Unlike the KM analysis, this latter approach is parametric and can infer the missing mutation status in the families thus yielding more precise penetrance estimates. As a result, the CIs are smaller under this approach, as shown in Table 3. Moreover, these CIs are in the same range than those published by another population-based study (Jenkins et al., Clinical Gastroenterology and Hepatology, 2006, 4:489-98) but are wider than those estimated in some clinic-based studies. The reason is that clinic-based studies are enriched with mutation carriers and although they provide more precise penetrance estimates, these estimates are likely to be biased and cannot be used for population inference. In some recent simulation studies that we have performed (Choi et al., Hum. Hered. 2008, 66:238-251), where we analyzed the tradeoff between bias and precision of penetrance estimates across various study designs, we showed that a population-based design with mutation carrier probands, as used in this paper, provided the best results (See our discussion). Despite this evidence, we agree that the results might still be too imprecise for the management of mutation carriers and this sentence was removed.

The last sentence of our conclusion (manuscript body and abstract) was replaced by " This study provides a unique population-based study of CRC risks among *MSH2/MLH1* mutation carriers in a Canadian population and can help to better define and understand the patterns of risks among members of Lynch Syndrome families."

Also I noted that patients with missense mutations are included in the study. Which could theoretically cause a lower penetrance.

Based on the in vitro functional assays and / or in silico prediction tools (e.g. SIFT, PolyPhen etc.) these missense substitutions were either predicted to be pathogenic or showed compromised MMR protein levels/function, so they were included in our analysis.

Furthermore probands with age of diagnosis of colorectal carcinoma under the age of 74 were included in the study. However hazard ratios are also mentioned in table 3 for the age group to 90 years.

These estimates are based on the probands' relatives but we agree that due to the small sample size they might not be very accurate, so we have changed the text and tables and now provide penetrance estimates until 70 years only.

In conclusion, I think the study design is good, but the authors should extend their group and publish the results then. I think if a sufficiently large sample is reached the results could very well be valuable.

Unfortunately, the prevalence of MMR mutations is very low in the general population and increasing the sample size of mutation carrier probands would require to screen a very large population of colorectal cancer patients which is difficult to achieve in Ontario.

Reviewer: Sining Chen

Below are specific edits and comments:

--Abstract: line -2 in Results: "The risk associated with ..." # "The relative risk associated with ..."

Done

--Intro: page 4 line -6: "type of mutation" # "mutated gene". Delete "the population of all".

Done

--Methods: the appendix explaining the logistic model, while important, is too short, not clear and reads like copy & paste from several other files. The authors did not explain the role of α and λ , in what way this is a logistic regression. The 1st page of the appendix also ended abruptly. It is also not clear why the authors introduced T and δ .

Further details have been added to appendix 1.

--Methods: page 11, para 2, line 2: “the probability of the observed diseases”#” the observed phenotypes”. It is also recommended that the ascertainment corrected likelihood be briefly explained in the appendix.

This change has been made and an appendix (#2) about the ascertainment correction has been added to the paper.

--Results: while the absolute risk results are plausible, the hazard ratio estimated from this method may not be applicable to the general population. This is because the estimated noncarrier risks are much higher than observed in the population, for example in table 3, a combined risk of 9% to age 70 and 23% to age 90. In the US population these are about 2% and 5% respectively. The authors should discuss the possible reasons and implications for the over-estimation and caution readers against using the estimated hazard ratios in clinical counseling. To that end, the values of the hazard ratios should be left out of the abstract.

The possible discrepancy between our cumulative risk estimates in non carriers vs. those published in the general US population might be due to the fact that our family sample is enriched with affected individuals that could have a different genetic and non-genetic risk profile than the general population. Therefore, even if our design and ascertainment correction approach tend to yield estimates that are closer to the general population, the difference seen between the two estimates can reflect a difference in the distribution of the genetic (besides the MMR mutations) and non-genetic risk factors in our sample compared to the general population (Begg, JNCI 2002;94: 1221-6). (Added to our discussion).