

## Reviewer's report

**Title:** Penetrance of HNPCC-related cancers in a retrospective cohort of 12 large Newfoundland families carrying a MSH2 founder mutation: an evaluation using modified segregation models

**Version:** 1 **Date:** 28 July 2009

**Reviewer:** Catherine Bonaïti-Pellié

### Reviewer's report:

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

- 1) Background, 3rd §: it should be stated more clearly that the ascertainment bias is essentially towards multiple cases families (the more severe outcomes is only a minor problem).
- 2) Background, last §: none of the quoted studies [15-18] mentioned any argument for an effect of lifestyle or environmental factors on penetrance. You should mention this possibility as a mere hypothesis and cancel the references.
- 3) Methods, Family ascertainment and characteristics, 2nd §: family members with HNPCC and a 50% of being a carrier should not be presumed to be carriers since some of these cancers are quite frequent in the general population, which is confirmed by the observation of cases among non-carriers: this assignment introduces a systematic error
- 4) Statistical methods: all the part regarding the Kaplan-Meier estimator should be removed. Indeed, there is no need to remake the statistical error of the paper of Green et al (2002). Contrarily to the affirmation of the authors, it does not allow "to assess potential bias ... mutation status" as the ML method performs much better. This analysis is even misleading, as physicians may be more familiar with this method than with a segregation-based method, and as most of them do not understand the problem of ascertainment bias, they would be tempted to believe in the results provided by K-M which are completely wrong (see Carayol et al., J Med Genet, 2002).
- 5) Discussion, page 6, last §: attributing the main difference between the results of this paper and those of Green et al. (2002) to the treatment of the relatives with unknown genotypes is an error, as it is mainly due to the absence of ascertainment correction (mentioned only at the end of this paragraph): again, see Carayol et al., 2002.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

- 1) Reference 7 is not relevant with what is written page 3, 1st column, 1st paragraph, lines 10-11.

2) Discussion: it is not correct to simply quote the different studies which provided risk estimates without any critical regard, which tends to confer the same degree of confidence to all the risk estimates. Indeed, some of them corrected for ascertainment bias, and some others did not, which can lead to considerable overestimate of risks, as stated above. Moreover, some more recent papers have been omitted: Alarcon et al. (Eur J Hum Genet, 2007) who corrected for ascertainment bias and Barrow et al. (Clin Genet, 2008) who tempted to correct for this bias but did not succeed in my view.

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

1) Methods: The ACR likelihood method is not sufficiently described in the appendix. As this is the main interest of this paper (the data have already been published), a complete description of the likelihood including how missing genotypes are handled, would be useful. On the other hand, I guess that correction for the ascertainment event includes a correction on the proband's genotype but this should be specified.

2) Results: page 5, 1st column, 1st §: In addition to the p-value of the best model, it would be interesting to indicate the p-value of the rejected model. In the following paragraph, the difference in CRC risk between male and female carriers is clear, but the other differences are not, with overlapping CIs: why not perform a test? Regarding the difference in risks between carriers and non-carriers, the result is not a scoop! Is it worth mentioning it? The same criticism regarding the absence of test applies to the 2nd § of column 2: in particular, the greater relative risk of developing CRC in female carriers could be a chance fluctuation.

3) Results, page 6, columns 1-2: Whereas, I am totally convinced of the interest of having correct estimates of risks for the clinical management of mutation carriers, I am not totally convinced by the illustration on the computation of risk of being a carrier. In the examples given, the risk for a relative of being a carrier computed with MMRpro is only slightly affected by a misspecification of penetrance: will it really modify the physician's attitude towards the counselee?

4) Discussion, page 7, 3rd §: the authors should also compare with the results of Alarcon et al (2007) who used the method of Carayol and Bonaïti-Pellié (2004), very close to the one used by Quehenberger et al. (2005).

5) Table 2, page 11: The cumulative CRC risks by age 70 appear surprisingly high: this would deserve a comment in the discussion.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a

statistician.

**Declaration of competing interests:**

I declare that I have no competing interests