LETTER TO JMG
Hereditary non-polyposis colorectal cancer: current risks of colorectal cancer largely overestimated

J Carayol, M Khlat, J Maccario, C Bonaiti-Pellié

A positive family history has been shown to be an important risk factor for colorectal cancer (CRC). Part of the familial aggregation is explained by the inherited diseases familial adenomatous polyposis and hereditary non-polyposis colorectal cancer (HNPCC). The latter syndrome is characterised by a high risk of colorectal cancer with a high rate of multiple primary tumours and a young age of onset, and also by a high risk of cancers of other organs (endometrium, stomach, pancreas, ovary, small intestine, urinary tract). The germline mutations which cause this syndrome have been shown to occur on genes that are responsible for repairing DNA mismatches. In humans, six mismatch repair (MMR) genes have been identified (hMLH1, hMSH2, hPMS1, hPMS2, hMSH6, and hMSH3) but germline mutations have been found in the first five only, mostly in hMLH1 and hMSH2. It is now commonly accepted that the lifetime risk of colorectal cancer in MMR carriers is very high, between 70% and 90%. A recent review of the available data indicated a lifetime risk of colorectal cancer of 74% or more in males, and a lifetime risk in females, and a lifetime endometrial cancer risk of 42% or more in female mutation carriers.

Apart from the preferential localisation of tumours in the proximal part of the colon, and the high frequency of multiple tumours, there is no specific individual characteristic of the syndrome. Therefore, the syndrome is diagnosed in patients on familial criteria. The classical criteria are the so-called “Amsterdam criteria,” which were issued in an effort to standardise clinical studies and to ensure that only families with HNPCC would be classified as such. These stringent criteria include: (1) three relatives with colon cancer, two of them being first degree relatives of the third; (2) at least two generations affected by colon cancer; and (3) one colon cancer patient diagnosed at 50 years or younger. New criteria for HNPCC were published after a workshop organised by the International Collaborative Group on HNPCC (Amsterdam II), which both substantially expanded the Amsterdam criteria to take into account extracolonic cancers and, in the Bethesda guidelines, early onset adenomas.

Because of the lack of DNA mismatch repair in the tumour cells of affected subjects, these cells acquire mutations that change the length of the nucleotide repeat sequences, termed microsatellite instability (MSI). This MSI can be detected in the tumours of affected subjects and Aaltonen et al. proposed using this phenotype as a prescreening method in HNPCC.

In most of the studies that have provided estimates of the risk for a carrier of one of the MMR genes of developing CRC or extracolonic cancer (ECC), the families were ascertained using the Amsterdam criteria. The lifetime risks for colorectal cancer ranged from 78% to 87%, slightly higher for males than females. The lifetime risks for extracolonic cancer also turned out to be substantial, with estimates ranging from 28% to 72%, the highest risk being for endometrial cancer in women, with values ranging from 43% to 60%. Yet the use of these very restrictive criteria is bound to cause an ascertainment bias towards multiple case families. The reason for this is that, in populations where sibships are small, families with no case, one case, or two cases, which are the majority among families including mutation carriers, do not meet the criteria for being selected. In addition, families with predominantly ECC cases will not be tested. Therefore, an over-representation of families with multiple CRC cases in the samples, and for this reason the cancer risk in mutation carriers, and in particular CRC risk, is expected to be overestimated in these studies.

The aim of the present paper is to investigate the bias associated with the Amsterdam criteria for both colorectal and extracolonic cancer risks. For this purpose, we simulated samples of at most three generation families, using the French population demographic characteristics and various values of the colorectal cancer and extracolonic cancer risks. We then tested subsets of HNPCC pedigrees meeting the Amsterdam criteria, and compared the cancer risk estimates based on these pedigrees to the actual risks underlying the simulations. Lastly, we discussed the importance of the bias in relation to the magnitude of the actual risks.

MATERIAL AND METHODS
The families which were simulated were at most three generation families, with two ancestors one of whom carried a mutation. Since the probability that a given family fulfils the Amsterdam criteria obviously depends on family size, we chose a simulation process where the family size and structure would be variable, as it is in reality, and we used French demographic data dating back to 1920 to perform the simulations. The ancestors in each family were taken from the generation of people born between 1901 and 1925. All the children of this couple and their grandchildren (third generation) were generated using the following parameters: number of children per woman and per birth cohort; interval between two consecutive births according to the mother’s birth cohort; age of the mother at the birth of the first child according to the mother’s birth cohort.

The simulation was conducted using the guidelines provided by Pennec et al. (1) each mother of the ancestral couple was randomly attributed a date of birth between 1901 and 1925, and (2) each mother in the pedigree was randomly attributed the total number of her children, her age at the birth of the first child, and the age at the following births, if any, according to her birth cohort. The descendants of a man were studied through those of his spouse, and the age difference between the spouses was arbitrarily set at two years.

The risks of cancer were assumed to be different for colorectal cancer and extracolonic cancer, and to vary with age and sex. Four age classes were considered: 20-39, 40-59, 60-79,

---

**Abbreviations:** CRC, colorectal cancer; HNPCC, hereditary non-polyposis colorectal cancer; MSI, microsatellite instability; ECC, extracolonic cancer
and 80-99 and the cancer risk was assumed to be nil before the age of 20. For each sex, and for each age class, a cumulative risk up to the end of the interval was fixed. The incidence rate \( \lambda_k \), being assumed to be constant over a given class \( k \), the cumulative risk \( RC(t) \) at age \( t \) belonging to the interval \( [t_k, t_{k+1}] \), may be obtained by the following formula:

\[
RC(t) = 1 - \exp \left[ - \sum_{i=1}^{k-1} \lambda_i (t_{i+1} - t_i) - \lambda_k (t - t_k + 1) \right]
\]

Two sets of risks defined as “low” and “high” risks were considered and are given in Table 1. The high risks are those given in previous studies and the low risks were arbitrarily chosen at values approximately 30% less than the high ones. For ECC, women were attributed a higher risk than men since endometrial and ovarian cancer are part of the spectrum of tumours.

In each pedigree, the genotypes were simulated according to Mendel’s laws for subjects whose parents were in the pedigree, as was the relative frequency of genotypes in the population for spouses. The frequency of the mutation in the population was arbitrarily fixed at 0.001 and we assumed the absence of de novo mutations. We considered the most informative situation in which all the genotypes of the subjects in the pedigree would be known and the phenotypes were simulated using the values of risks given in Table 1. For each person, at each age greater than 20 years and until age at last contact, the disease status was randomly attributed for CCR as well as ECC. In order to avoid the difficult problem of independence of cancers in case of multiple tumours, the first occurrence of cancer was the only one to be considered. Thus, any person affected at a given age with a type of cancer was censored at this age for other events. For the sake of simplicity, we did not simulate death and the age at last contact of all subjects was the age they would have reached in the year 2000.

We also considered that families with many affected subjects have a greater likelihood of being ascertainment biased than families with smaller numbers. If \( \pi \) is the probability that a family fulfilling the Amsterdam criteria is ascertainment biased, this probability was set to 1 in a first step, and then we considered the two situations where \( \pi \) would be equal respectively to N/6 and N/15, N being the number of CCR cases in the pedigree.

For each model of risks, 40 replicates of 1000 pedigrees were simulated and the simulation procedure was validated by checking that the risk estimates based on the whole sample were the same as the theoretical risks underlying the simulation.

The risks were estimated using the Kaplan-Meier estimator, as used by all authors. The bias was evaluated through the average difference between the actual risk and the estimated risk:

\[
B = \frac{1}{n} \sum_{i=1}^{n} (R_i - R_0)
\]

where \( R_i \) is the risk estimated in the replicate \( i \), and \( R_0 \) is the actual risk used in the simulation.

**RESULTS**

Whatever the risks, the number of subjects aged more than 80 was too small to permit estimation of cancer risk in this age class and therefore the results are given for the first three classes only.

| Table 1 Cumulative risks considered in the simulations for colorectal cancer (CRC) and extracolonic cancers (ECC) according to sex and age |
|---|---|---|---|---|
| **CRC risk** | **ECC risk** |
| **High** | **Low** | **High** | **Low** |
| **Male** | **Female** | **Male** | **Female** |
| 39 years | 59 years | 79 years | 99 years |
| 0.20 | 0.20 | 0.15 | 0.10 |
| 0.50 | 0.30 | 0.65 | 0.35 |
| 0.80 | 0.60 | 0.70 | 0.50 |
| 0.90 | 0.70 | 0.50 | 0.60 |

| Table 2 Cumulative colorectal cancer (CRC) risk estimates and biases when applying Amsterdam criteria on 40 replicates of 1000 families carrying a HNPCC mutation, according to the actual risk in the simulated sample |
|---|---|---|---|
| **Actual CRC risk** | **Estimated CRC risk** | **Bias (%)** | **Range (min–max)** |
| **“High” risks** |
| **Men** | 39 | 0.2 | 0.29 | 46 | 0.21–0.34 |
| 59 | 0.5 | 0.68 | 37 | 0.60–0.75 |
| 79 | 0.8 | 0.93 | 17 | 0.89–1.00 |
| **Women** | 39 | 0.2 | 0.31 | 56 | 0.23–0.37 |
| 59 | 0.4 | 0.59 | 48 | 0.51–0.65 |
| 79 | 0.7 | 0.88 | 26 | 0.81–0.96 |
| **“Low” risks** |
| **Men** | 39 | 0.1 | 0.21 | 110 | 0.12–0.34 |
| 59 | 0.3 | 0.57 | 90 | 0.45–0.68 |
| 79 | 0.6 | 0.87 | 45 | 0.64–1.00 |
| **Women** | 39 | 0.1 | 0.23 | 130 | 0.16–0.36 |
| 59 | 0.2 | 0.45 | 125 | 0.32–0.68 |
| 79 | 0.5 | 0.81 | 62 | 0.62–1.00 |

Applying a selection according to the Amsterdam criteria (with complete ascertainment of families fulfilling these criteria) dramatically reduced the sample size, as expected. Indeed, only about 10% of the families in the samples fulfilled these criteria, when the actual CRC risks were high, as defined in Table 1, and only 4% when these risks were low. In our analysis, the estimates of CRC and ECC risks are independent given that we studied only the first occurrence of cancer, and the biases associated with each of these risks are shown in Tables 2 and 3. For the CRC risk estimates, there is a very large overestimation of risks from 17% to 130% depending on the actual risks. Compared to the CRC risks, the ECC estimates are substantially underestimated (about 30% lower than the actual values), with a small variation according to the actual risks. Because of random fluctuations, the range of estimates is generally large, for both cancer risks, as shown in Tables 2 and 3.

On one hand, selecting families on Amsterdam criteria and not taking into account this selection in the estimation method largely overestimates the CRC risks. On the other hand, there is also an underestimation of the ECC risks and that was quite unexpected; indeed, since families are selected...
through CRC cases only, the proportion of other cancers in families should remain unchanged. However, since subjects affected with CRC are censored at the age of onset of this cancer for other events, selecting CRC cases as a first occurrence results in selection of cases not affected with ECC. To get an idea of what the actual CCR and ECC risks could be, we simulated samples using various values of risks, until we reached the CCR and ECC risk estimates closest to the average estimates which were found in published studies (table 4). We found that the actual CRC lifetime risks could be as low as 40%, instead of 80% in men and 70% in women, and that the ECC risks could be as high as 60% in men, instead of 35%, and 65% in women, instead of 50%.

Allowing for preferential ascertainment of families with multiple cases of CRC did not modify the estimates of risks (results not shown), whatever the value of \( \pi \) specified in the simulation. However, since it drastically reduced the sample size, the random fluctuations were substantially greater.

**DISCUSSION**

Our study clearly showed that colorectal cancer risks are largely overestimated in HNPPC, at about double the actual levels. It also showed a not negligible underestimation of the risk of extracolonic cancer, which would be still higher than CRC risks. In summary, the lifetime risk of colorectal cancer for carriers would be about 40% in both sexes, and the lifetime risk of extracolonic cancer would be about 60% in men and 65% in women.

In most simulation studies dealing with ascertainment biases, families have a fixed size, usually relatively small (about three sibs per sibship). The problem with this constraint is that families meeting the Amsterdam criteria would have very few unaffected subjects, and therefore also very few unaffected carriers, which would artificially inflate the bias on the risk estimates. To avoid this problem, we have developed a simulation method using real demographic data. In our simulated samples, the families meeting the Amsterdam criteria are larger than the others, allowing for a proper representation of unaffected carriers. However, since we used French demographic parameters to perform the simulations, one could argue that these parameters might not be appropriate to the countries (The Netherlands, Finland, USA) in which the risks have been estimated. To evaluate the sensitivity of our results to this parameter specification, we performed the simulations and evaluated the biases in very extreme, and unrealistic, situations where the fertility rate of women would be twice or half the French one. As expected, the bias was higher for a low fertility rate and lower with a high one, but the impact on the risk estimates was quite small. For instance, for an actual cumulative risk of 60% of developing colorectal cancer at the age of 79, which would lead to an estimation of 87% (bias 45%) using French fertility rates, the estimated risk would be 79% (bias 32%) if the fertility rate was twice, and 94% (bias 57%) if this rate was half the French one. Given that the countries where the studies were carried out have demographic parameters which are fairly comparable to the French one, we can conclude that the biases found in this study are quite robust to demographic parameter specification.

Would these results be different in a more realistic situation where family members would not all be tested, some of them could be diseased, de novo mutations could occur, etc.? Assuming mortality rates in the simulation process would decrease the overall information provided by a pedigree. Indeed, some subjects would be studied at various ages and a smaller number of subjects would be available for risk estimation, particularly in the older age groups. However, this is not likely to affect the risk estimates, given that mortality from other diseases, according to French demographic data. Therefore, an allowance for mortality would not have modified the cancer risk estimates, but rather would have reduced the precision of those estimates.

Unavailability of some family members is also likely to lessen the information provided by the pedigrees. In “real life”, a systematic bias could be introduced if genetic testing would be independent of phenotype. For instance, affected subjects may be dead and unaffected subjects may more or less be willing to undergo genetic testing. An additional overestimation of risk might be expected if affected relatives were
systematically considered as carriers (since some of them could be sporadic) and/or if a not negligible proportion of unaffected relatives were missing, which is probably the case in most of the studies which we referred to.

De novo mutations were not considered in our study. Such cases are unlikely to be ascertained using Amsterdam criteria since their family history is expected to be negative. Even when subjects are ascertained through non-familial criteria (as further discussed), patients’ parents are seldom available and de novo mutations are very difficult to prove directly. The argument for the existence of such mutations is usually indirect, through the study of haplotype sharing among subjects carrying the same germline mutation. This phenomenon is not expected to have any impact on our conclusions.

Among the hypotheses made in the present study, the most questionable is the assumption of genetic homogeneity, that is, of equal risk associated with all mutations. Previous studies showed that the risks associated with hMLH1 and hMSLS2 mutations appeared quite similar or only slightly different. However, we cannot exclude that some specific mutations could be associated with much higher risks than others. The high aggregation of colorectal cancer cases in Finnish families is striking, and it is possible that mutation 1 of hMLH1, a predominant mutation owing to a strong founder effect in Finland, is associated with particularly high colorectal cancer risks. If heterogeneity existed, we would expect that the estimation bias would be smaller for these families who would be over-represented in a sample selected on Amsterdam criteria.

We would like to emphasise that the overestimation of risks shown in this study is a totally different issue from the discrepancies in penetrance estimates found in other family syndromes, depending on the population in which the mutation carriers have been ascertained. Such discrepancies have been shown, for instance, in breast-ovarian cancer family syndrome resulting from mutations of the BRCA1 or BRCA2 genes. The estimation from multiple case families, although using a method which perfectly corrected for ascertainment bias, provided higher risk estimates for BRCA1 mutation carriers than studies of subjects selected independently of family history in a population with a particularly high frequency of a specific BRCA1 mutation. Such a discrepancy could be explained by the fact that the cancer risk conferred by mutated BRCA1 is modified by other factors, either genetic or acquired, that themselves run in families.

The question is now how could unbiased estimates of cancer risks be obtained? The risks that we found by our approximate method give an order of magnitude. Dunlop et al proposed selecting subjects on age at diagnosis of the index case (at or below 35 years) and the presence of MSI in tumours.

• The lifetime risk of colorectal cancer among mutation carriers in HNPCC families is considered to be very high (70-90% in published reports).

• This risk is likely to be underestimated as it is based on families selected according to the very stringent Amsterdam criteria and not corrected for ascertainment bias.

• Using a simulation study, we showed that the bias was such that the estimated risk could be double the actual risk.

ACKNOWLEDGEMENTS

This work was supported in part by the French “Ligue Nationale Contre le Cancer”.

Authors’ affiliations

J Corayol, C Bonaiti-Pellié, Unité de Recherches en Épidémiologie des Cancers (INSERM U521), Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif Cedex, France

M Khlat, Unité de Recherches en Épidémiologie des Cancers (INSERM U535), Bâtiment Grégoire Pincus, CHU du Kremlin-Bicêtre, 80 rue du Général Leclerc, 94276 Le Kremlin-Bicêtre Cedex, France

J Maccario, Unité de Recherches en Épidémiologie (INSERM U472), Paris, France

Correspondence to: Dr C Bonaiti-Pellié, INSERM U525, Bâtiment Grégoire Pincus, CHU du Kremlin-Bicêtre, 80 rue du Général Leclerc, 94276 Le Kremlin-Bicêtre Cedex, France; bonaiti@kb.inserm.fr

REFERENCES


6 Rodríguez-Bigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan FM, Lynch H, Perucch M, Smyrk T, Sabon L, Snivstova S. A National Cancer Institute Workshop on Hereditary Nonpolyposis...
TNF2 turns up in non-ischaemic heart failure

Tumour necrosis factor α (TNF α) gene polymorphism and non-ischaemic heart failure are linked, according to latest research by Densem et al, with the “TNF2” allele predominating in patients with this type of heart disease.

TNF α overexpression is linked with heart failure and its effects are reversible with TNF α binding protein in animal studies. TNF α concentrations in blood are raised in patients with heart failure and are proportional to severity of symptoms and survival. Early trials indicate that TNF α inhibition decreases symptoms. Healthy people vary in the amount of TNF α they produce—according to whether they have a G to A substitution at position—308 in the gene promoter. The TNF2 allele results in a six- or sevenfold increase in inducible synthesis.

Densem et al reasoned that frequency of the TNF2 genotype would be higher in patients who had a heart transplant for end stage heart failure than in the general population and compared expected against actual frequency in 175 transplant patients and 212 unrelated healthy controls. TNF2 allele frequency was similar for the transplant patients and controls (31% v 27%). However, it was significantly higher in non-ischaemic (viral/idiopathic) than ischaemic heart failure (38% v 27%, p=0.03) and the controls (38% v 27%, p=0.05) when patients were grouped according to their diagnosis before transplantation. TNF2 frequency was similar in patients with ischaemic heart failure and controls.

If TNF α is a cause of non-ischaemic heart failure—currently unproved—anti-TNF α treatment would be beneficial. 

† Heart 2002;87:153–155.