Risk of colorectal and endometrial cancer for carriers of mutations of the hMLH1 and hMSH2 gene: correction for ascertainment

F Quehenberger, H F A Vasen, H C van Houwelingen

Background: Hereditary non-polyposis colorectal cancer (HNPCC) is caused by germline mutations of mismatch repair genes, usually in hMLH1 or hMSH2. All earlier studies on penetrance except one population based study were conducted in HNPCC families and did not correct for the way in which these families were ascertained.

Objective: To obtain estimates of the risk of colorectal cancer (CRC) and endometrial cancer (EC) for carriers of disease causing mutations of the hMSH2 and hMLH1 genes.

Methods: Families with known germline mutations of hMLH1 (n = 39) and hMSH2 (n = 45) were extracted from the Dutch HNPCC cancer registry. Ascertainment-corrected maximum likelihood estimation was carried out on a competing risks model for cancer of the colorectum and endometrium.

Results: Both loci were analysed jointly as there was no significant difference in risk (p = 0.08). At age 70, colorectal cancer risk for men was 26.7% (95% confidence interval, 12.6% to 51.0%) and for women, 22.4% (10.6% to 43.8%); the risk for endometrial cancer was 31.5% (11.1% to 70.3%).

Conclusions: Current estimates of the CRC risk of mutations to the hMLH1 and hMSH2 locus should be replaced by considerably lower risks which account for the selection of the families.

Hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome[1] was originally defined as familial clustering of colorectal cancer (CRC). A set of diagnostic criteria, the so called Amsterdam criteria, was proposed to provide uniformity in clinical studies. According to these criteria, at least three relatives should have colorectal cancer (that is, Amsterdam criteria I) or HNPCC associated cancers (cancer of colorectum, endometrium, ureter, renal pelvis, and small bowel (Amsterdam criteria II)), and one of them should be a first degree relative of the other two. In addition, at least two generations should be affected, the cancer should be diagnosed before the age of 50 years in one of the relatives, familial adenomatous polyposis should be excluded, and the cancer should be confirmed by pathology.[2]

HNPCC is caused by germline mutations in DNA mismatch repair genes (MMR). Mutations of hMSH2 and hMLH1 constitute almost 90% of the mutations reported in families with HNPCC and are identified in half the families which meet the Amsterdam criteria.[3] The penetrance of these mutations was found to be high,[1] but the evidence was mainly based on observed risk of CRC in cohorts of mutation carriers that were identified by multiple cancer cases in the families. This contains a circular argument, as the excess of CRC cases—which was the reason for genotyping family members—was counted again for a risk estimate. On the other hand, a family with only two cancer cases and an equal

number of mutation carriers, which would provide evidence for low mutation risk, would not be included in the study. Indeed it was shown that this method seriously over-estimated the CRC risks of mutations and underestimated the risks of extracolonic cancers in families that were ascertained by the Amsterdam criteria I.[4]

Direct assessment of penetrance by identifying cohorts of mutation carriers from the population would be difficult, as the frequency of hMSH2 and hMLH1 mutations in CRC cases was found to be as low as 0.135% by Salovaara *et al*[5] and 0.0319% by Dunlop *et al*.[6] In this situation, the genotyping of a kin cohort of relatives of cancer cases is a possible way of identifying more mutation carriers. This design was later described as genotyped proband or kin cohort design, with or without additional genotyping of relatives.[7–9]

The most efficient way of identifying mutation carriers is to ascertain families with multiple cancer cases. If a mutation has been identified in such a family, genetic testing is offered to all unaffected relatives. The proportion of carriers of the pathogenic mutation among unaffected relatives is expected to decrease with age because they have developed cancer, whereas affected relatives not carrying the pathogenic mutation are increasingly found at older ages. From these observations it is possible to draw conclusions about the cancer risk associated with the mutation, independent of the phenotypic criteria that were originally used to ascertain the high risk families.[10] [11]

Our aim in this study was to obtain estimates of the risk of CRC and endometrial cancer (EC) for carriers of disease causing mutations of the hMSH2 and hMLH1 genes from data on HNPCC families in which a mutation of either the hMSH2 or hMLH1 gene had been found by conditioning on all phenotypic information.

METHODS
HNPCC families

In 1987 a national registry for families with HNPCC was established in the Netherlands. The registry had three objectives:

- to promote surveillance in HNPCC families;
- to safeguard the continuity of the surveillance programme;
- to promote research.

The methods and approach of the registry have been described elsewhere.[12] In brief, clinical specialists or clinical genetic centres from all parts of the Netherlands refer all families suspected of HNPCC because of clustering of CRC to the registry. The genealogical studies were carried out by genetic field workers associated with the registry or by

. . . .

clinical geneticists. We collected clinical information including the age at diagnosis of cancer, site of the tumour, age at death, and causes of death. The cancer diagnosis was confirmed by medical and pathological reports in as many affected relatives as possible. In addition, we collected data on colonoscopic screening of the unaffected relatives. For the present study families were selected in which a disease causing mutation of hMLH1 or hMSH2 had been identified. Genetic counselling and testing were offered to all first degree relatives of carriers of a pathogenic mutation.

Mutation analysis

The techniques used in the mutation analysis have been reported previously.[13] To summarise, the general strategy was to amplify by polymerase chain reaction each of the 16 exons of hMSH2 and the 19 exons of hMLH1 in a single affected member of the family, and to analyse these products by guanosine and cytidine extension clamped denaturing gradient gel electrophoresis. To determine the molecular nature of the variant, exons with an altered pattern of migration on denaturing gradient gel electrophoresis were sequenced. When variants were detected, the investigations were extended to the rest of the family to verify the segregation of the nucleotide change with the disease phenotype.

Study data

The identity of the parents was recorded for each non-founder family member. The phenotypic data included the current age, age at first colonoscopy, age at death, age at diagnosis of CRC, EC, or cancer at minor HNPCC sites (MC). The minor HNPCC sites were the small bowel, the stomach, the ovary, and the urinary tract including the renal pelvis and ureter but excluding the urinary bladder. Mutation status at

Table 2 Parameter estimates of the final model

Cancer site	Polynomial	Coefficient	Estimate (SE)
CRC	$a_0 + a_1 t + a_2 t^2$	a_0	2.7 (0.46)
		a_1	−2.0 (0.48)
		a_2	−1.2 (0.43)
EC	b_0 (constant)	b_0	3.4 (0.60)
MC	c_0 (constant)	c_0	1.7 (0.47)

Estimated coefficients and standard errors of polynomial functions of degree 2 and 0, respectively, for the logarithm of the cause specific relative hazard rates for carriers of a mutation at the either hMLH1 or hMSH2 locus relative to the population.
CRC, colorectal carcinoma; EC, endometrial carcinoma; MC, minor HNPCC site (small bowel, stomach, ovary, and urinary tract including the renal pelvis and ureter but excluding the urinary bladder); t, time scale derived from age in years by subtraction of 55 and division by 20.

the hMLH1 and the hMSH2 locus was obtained from a part of the family members.

Statistical analysis

HNPCC-causing germline mutations of MMR genes were assumed to have a population frequency of 0.1% each,[5][6] to follow Mendelian inheritance, and to be in Hardy–Weinberg equilibrium. All calculations were conditional on the family structure.

A biallelic single locus model was applied, which meant that in hMLH1 families only the hMLH1 locus was modelled and in hMSH2 families only the hMSH2 locus was modelled. The unconditional likelihood of observing the genotypes and the phenotypes of a pedigree was factorised into the product of the likelihood of the genotypes and the likelihood of the phenotypes conditional on the genotypes. The likelihood of the genotype was determined by the mutation allele

Table 1 Characteristics of HNPCC families in which either an hMLH1 or an hMLH2 mutation had been found

Variable	hMLH1		HMSH2	
Number of pedigrees	39		45	
Subjects in the largest pedigree	108		70	
Generations in the deepest pedigree	6		5	
Amsterdam criteria I (% of total)	23	(59%)	32	(71%)
Amsterdam criteria II (% of total)	29	(74%)	40	(89%)
Mean size of pedigrees (median)	27.9	(24)	29	(25)
Number of subjects with phenotype information	1088		1304	
Non-diseased subjects				
Number of subjects (% of total)	846	(77.8%)	1028	(78.8%)
Genotyped (% of non-diseased)	319	(37.7%)	363	(35.3%)
Mutation carriers (% of genotyped)	156	(48.9%)	149	(41.0%)
Diseased subjects (CRC, EC, MC)				
Number of subjects (% of total)	242	(22.2%)	276	(21.2%)
CRC (% of diseased)	202	(83.5%)	207	(75.0%)
EC (% of diseased)	27	(11.2%)	36	(13.0%)
MC (% of diseased)	13	(5.4%)	33	(12.0%)
Genotyped (% of diseased)	45	(18.6%)	51	(18.5%)
Mutation carriers (% of genotyped)	44	(97.8%)	48	(94.1%)
CRC (% of mutation carriers)	37	(84.1%)	32	(66.7%)
EC (% of mutation carriers)	4	(11.4%)	11	(22.9%)
MC (% of mutation carriers)	2	(4.5%)	5	(10.4%)
CRC before EC or MC	9		17	
EC before CRC or MC	13		14	
Mean age at first cancer				
CRC men	43.4		45.5	
CRC women	42.3		43.4	
EC	45.7		43.3	
MC	50.1		48.6	

CRC, colorectal carcinoma; EC, endometrial carcinoma; HNPCC, hereditary non-polyposis colorectal cancer; MC, cancer at minor HNPCC sites.
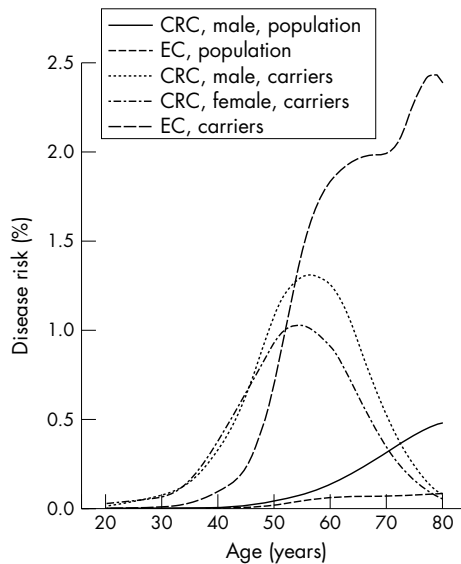
Figure 1 Age dependent incidences of the population and of carriers of a pathogenic mutation at either hMLH1 or hMSH2 locus. CRC, colorectal carcinoma; EC, endometrial carcinoma.

frequencies. For the likelihood of the phenotypes conditional on the genotypes we assumed an age dependent disease risk for carriers and non-carriers of a pathogenic mutation. CRC, EC, and MC were determined to be competing risks.[14] This meant that any information about cancer cases was ignored after the first diagnosis of any of CRC, EC, or MC. Assuming that the minimum of current age, age at first colonoscopy, and age at death constitute an uninformative censoring event, the partial likelihood of observing a non-diseased person with genotype g at the censoring time t was

$$S_{CRC,g}(t) \bullet S_{EC,g}(t) \bullet S_{MC,g}(t)$$

where S denotes the cause specific survival function, which is the exponential function of minus the cumulative cause specific hazard function. The complement of this function is the cause specific cumulative risk function; it gives the age dependent probability of getting the specific cancer diagnosed if there were no censoring and the risk of observing an event was not changed by the diagnosis of a competing cancer type. For persons diagnosed with the first cancer of type CRC, EC, or MC at age t, the likelihood of the phenotype given the genotype was the specific hazard function of that type of cancer multiplied by $S_{CRC,g}(t) \bullet S_{EC,g}(t) \bullet S_{MC,g}(t)$. In order to correct for ascertainment, we conditioned the likelihood of the observed genotypes and phenotypes on the likelihood of the observed phenotypes and on the event that at least one CRC, EC, or MC case in the family was a mutation carrier, which was given by the ratio of the two likelihoods.[10 15] The likelihood of the conditioning event was calculated as the difference between the likelihood of the phenotypes alone and the likelihood of the phenotypes and none of the cancer cases being a mutation carrier. More details on the statistical methods are given in the electronic appendix (this can be viewed on the JMG web site: www.jmedgenet.com/supplemental).

For each sex, cancer specific hazard rates for non-carriers were taken from the age dependent cancer incidences published by the Netherlands Cancer Registry for each five year age interval.[16] Thereby, we assumed that the mutation frequencies were so low that they would not have a substantial influence at the population level. Furthermore, we assumed that age dependent cancer incidences of the populations represented a cause specific hazard function.[14]

Finally the cause specific hazard functions were smoothed by a triangular kernel smoother with a kernel width of 11 years (see electronic appendix).

Cancer specific hazard rates of mutation carriers were modelled as a product of non-carrier hazard rates and an age dependent relative hazard function. We assumed that the logarithms of the cancer specific relative hazard rates were polynomial functions. The relative hazard rates of MC were set constant. The polynomial degrees for CRC and ED were determined by a two stage search. In the first stage a backwards search was made, reducing the polynomial degree for both cancers simultaneously. In the second stage a stepwise search was carried out separately for each cancer until there was no indication of a lack of fit according to the likelihood ratio criterion at the 5% level. This search was done for hMLH1 and hMSH2 families separately. For the comparison of cancer specific log relative hazard between hMLH1 and hMSH2 we added to the model relative hazard parameters for locus, one parameter for each cancer. Sex differences were tested similarly by adding a risk parameter.

Parameter estimates were obtained by maximum likelihood. Standard errors were obtained from the information matrix of the parameters by the delta method. Confidence intervals of disease risks were calculated symmetrically around the logarithm of the disease specific relative hazard rates. Hypothesis tests were based on the likelihood ratio criterion at a 5% error rate.

All the analyses were carried out on a modified version of the computer program MENDEL 3.3.[17]

RESULTS

The characteristics of the HNPCC families are summarised in table 1. The diseased non-carriers were one patient with ovarian cancer diagnosed at age 55 from an hMLH1 family; and two patients with CRC diagnosed at ages 73 and 42, respectively, plus one patient with gastric cancer diagnosed at age 28 from separate hMSH2 families. We assumed a constant relative hazard function for MC risk, as there were only seven cases of MC that had been genotyped. The search for the best fitting relative hazard function for CRC risk found second degree polynomials for both genes. The degree of the polynomial for the log relative hazard function for EC was zero, indicating that the relative hazard for EC did not vary with age. Sensitivity analysis showed little dependence of the results on the assumed allele frequencies, non-carrier risks, and phenotypic and genotypic information earlier than 1990 (table 5 in the electronic appendix).

The risk of hMSH2 mutations relative to hMLH1 was 0.23 (95% confidence interval, 0.03 to 1.80) in CRC and 0.47 (0.22 to 1.10) in EC. As the statistical test for differences between genes was not significant (p = 0.08), it was decided to assume the same hazard rates for both genes.

The test for a sex specific relative hazard rates was not significant for either CRC (p = 0.99) or MC (p = 0.88). However, we still assumed sex dependent hazard rates for non-carriers, as there was no evidence against this.

The coefficients of the polynomials that represented the logarithms of the relative hazard functions are shown in table 2. The resulting age dependent relative and absolute hazards are given in table 3 and fig 1. Maximum relative hazard for CRC was (mean (SD)) 33.1 (15.2) at age 39 years. The maximum age dependent incidence, which is given approximately by the hazard rates, was 1.3 (0.6)% at age 57 for men and 1.0 (0.5)% at age 55 for women. At age 72, CRC incidence had decreased to the population level. The age dependent incidence followed the population incidence at a distance, rising steeply after the age of 50 to a maximum at 80 years, as the relative risk of EC had not been found to be age dependent.

Table 3 Cause specific hazards relative to the population and absolute cause specific hazards of carriers of mutations at either the hMLH1 or the hMSH2 locus

	Age (years)	Relative risk*		Absolute risk (%)*	
CRC (male)	30	25.6	(15.2 to 43.0)	0.073	(0.044 to 0.12)
	40	33	(21.4 to 51.0)	0.33	(0.21 to 0.51)
	50	22.7	(14.5 to 35.4)	1.1	(0.70 to 1.7)
	60	8.3	(5.3 to 12.9)	1.3	(0.82 to 2.0)
	70	1.6	(0.92 to 2.8)	0.54	(0.31 to 0.95)
	80	0.17	(0.064 to 0.44)	0.081	(0.031 to 0.21)
CRC (female)	30	25.6	(15.2 to 43.1)	0.068	(0.040 to 0.11)
	40	33	(21.4 to 51.0)	0.38	(0.25 to 0.59)
	50	22.7	(14.5 to 35.4)	0.94	(0.60 to 1.5)
	60	8.3	(5.3 to 12.9)	0.91	(0.59 to 1.4)
	70	1.6	(0.92 to 2.8)	0.34	(0.20 to 0.60)
	80	0.17	(0.064 to 0.44)	0.058	(0.022 to 0.15)
EC	30	29.1	(16.5 to 51.1)	0.009	(0.0051 to 0.016)
	40	29.1	(16.5 to 51.1)	0.096	(0.055 to 0.17)
	50	29.1	(16.5 to 51.1)	0.72	(0.41 to 1.3)
	60	29.1	(16.5 to 51.1)	1.9	(1.0 to 3.3)
	70	29.1	(16.5 to 51.1)	2	(1.1 to 3.5)
	80	29.1	(16.5 to 51.1)	2.4	(1.4 to 4.3)
MC (male)	30	5.87	(3.62 to 9.5)	0.011	(0.0068 to 0.018)
	40	5.87	(3.62 to 9.5)	0.052	(0.032 to 0.084)
	50	5.87	(3.62 to 9.5)	0.24	(0.15 to 0.38)
	60	5.87	(3.62 to 9.5)	0.68	(0.42 to 1.1)
	70	5.87	(3.62 to 9.5)	1.7	(1.0 to 2.7)
	80	5.87	(3.62 to 9.5)	2.5	(1.5 to 4.0)
MC (female)	30	5.87	(3.62 to 9.5)	0.034	(0.021 to 0.056)
	40	5.87	(3.62 to 9.5)	0.1	(0.064 to 0.17)
	50	5.87	(3.62 to 9.5)	0.27	(0.17 to 0.44)
	60	5.87	(3.62 to 9.5)	0.54	(0.34 to 0.88)
	70	5.87	(3.62 to 9.5)	0.91	(0.56 to 1.58)
	80	5.87	(3.62 to 9.5)	1.2	(0.73 to 1.90)

The logarithms of cause specific relative hazards for colorectal carcinoma (CRC) were assumed to form a quadratic polynomial, which was the same for both sexes. The relative cause specific hazards for endometrial cancer (EC) and minor HNPCC cancer site (MC) were assumed to be constant.
Hazard rates are given as percentages, as they correspond approximately to age dependent incidence rates.
*95% confidence intervals are given in brackets.

The mutation carrier cumulative risks are presented in table 4 and fig 2. The results of separate analysis of hMLH1 and hMSH2 families are given in tables 2a to 4a and tables 2b to 4b, respectively, in the electronic appendix.

DISCUSSION
By assessing the occurrence of germline mutations of MMR genes in high risk families for a person aged 70 years carrying an hMLH1 or an hMSH2 mutation we found a CRC risk of 26.7% for men and 22.4% for women, whereas the risk for EC was estimated to be 31.5%. The highest incidence for CRC was 1.3% at age 57 for men and 1.0% at age 55 for women. The difference between genes was not significant. There was also no sex difference in relative hazard rates for CRC for mutation carriers.

Cancer risks of mutation carriers within HNPCC families have been investigated in at least seven studies.[13 18–23] The cumulative CRC risk reported for hMLH1 or hMSH2 ranged from 54%[13] to 100%[20] at age 70. Differences between genes[13 24] and between sexes[9] have been found in some studies. EC risks between 24% and 62% have been reported.[13 20–23] All studies except one[9] applied the Kaplan–Meier estimate to cohorts of mutation carriers that were ascertained on the occurrence of multiple cancer cases within a family. Carayol et al[4] had demonstrated in a simulation study that by selecting HNPCC families under the demographic parameters of France by the Amsterdam criteria I the CRC risk estimates reported in previous studies were largely overestimated. They found that an actual male cumulative risk for CRC of 25% at age 59 would on average give a risk of 59% with the Kaplan–Meier estimate. Vasen et al,[13] using an earlier version of the

data which were underlying this study, found 55% and 70% for the risk of hMLH1 and hMSH2, respectively, at age 70. By analysing our data in a way that took into account that the families were selected on familial clustering of cancers and the occurrence of mutation carriers,[10 15] we found cumulative risks that were in close agreement with those obtained in the simulation study[4] (table 4).

Carayol et al[4] found in addition an underestimation of extracolonic cancer risk for the Kaplan–Meier method in families ascertained only by multiple CRC cases. At least some of the families in the current study were ascertained for extracolonic cancers, so we would expect a risk estimate for EC or MC, respectively, within the range reported by other studies. Unfortunately, owing to the few observed EC cases, our results are too imprecise to allow conclusions to be drawn.

Our analysis urged us to be explicit about the phenotype on which the families were selected and how phenotypes depended on genotypes. The relevant phenotypes were age at diagnosis of HNPCC specific cancers, at the start of screening, at death, or at study end. The competing risks model for the phenotypic data had the advantage that it took into account only the first HNPCC specific cancer and therefore did not need to be explicit about the risk in persons in whom one cancer had occurred already. The method is not sensitive to the selection of probands for genotyping, as it is unbiased under any sampling scheme.

The usual risk estimates of an event from competing risk models give the probability of observing the index event (or several index events) as the first of a set of events.[14] However, the risk estimates given in table 4 assume that the subject is
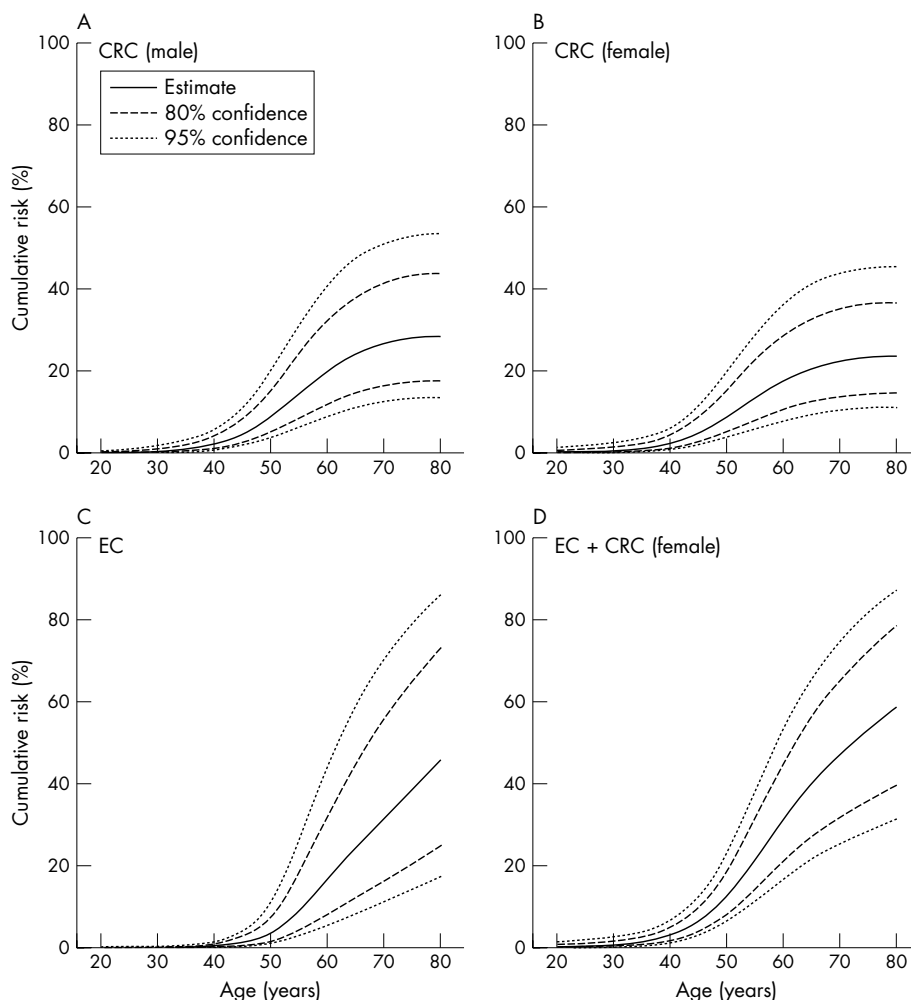
Figure 2 Age dependent cause specific cumulative risk and confidence intervals for carriers of mutations at either the hMLH1 or the hMSH2 locus. (A) Colorectal carcinoma (CRC) for men. (B) CRC for women. (C) Endometrial carcinoma (EC). (D) Both CRC and EC. It was assumed that no competing risks took effect.

still at risk for the index event after an event from the remaining set of competing risks has occurred. Nevertheless these estimates are different from an estimate of the risk of the index event in a population based cohort study. The reason for this is that by ignoring any events after the first one, competing risk models also ignore the possibly higher risk of a person getting the index cancer who had been diagnosed as having another HNPCC related cancer, compared

Table 4 Age dependent cause specific cumulative risks and 95% confidence intervals for carriers of mutations at either the hMLH1 or the hMSH2 locus

	CRC (male)		CRC (female)		EC (female)		CRC+EC (female)	
Age (years)	Risk	95% CI	Risk	95% CI	Risk	95% CI	Risk	95% CI
30	0.49	(0.13 to 1.8)	0.63	(0.14 to 2.8)	0.04	(0.01 to 0.13)	0.67	(0.15 to 2.9)
40	2.3	(0.90 to 5.7)	2.5	(1.0 to 6.5)	0.48	(0.15 to 1.5)	3	(1.3 to 7.0)
50	8.9	(3.9 to 19.9)	9	(3.9 to 19.9)	3.7	(1.1 to 11.3)	12.3	(6.3 to 23.3)
60	19.8	(9.0 to 40.5)	17.6	(8.0 to 36.2)	16.6	(5.5 to 44.1)	31.3	(16.8 to 53.5)
70	26.7	(12.6 to 51.0)	22.4	(10.6 to 43.8)	31.5	(11.1 to 70.3)	46.9	(25.1 to 75.1)
80	28.5	(13.7 to 53.4)	23.7	(11.3 to 45.5)	45.6	(17.3 to 85.8)	58.5	(30.9 to 87.6)
	MC (male)		MC (female)		CRC+EC+MC (male)		CRC+EC+MC (female)	
30	0.064	(0.023 to 0.17)	0.36	(0.13 to 0.97)	0.53	(0.16 to 1.8)	1	(0.37 to 2.7)
40	0.37	(0.14 to 0.98)	1	(0.38 to 2.7)	2.6	(1.1 to 6.0)	3.9	(2.0 to 7.6)
50	1.6	(0.61 to 4.4)	2.8	(1.0 to 7.5)	10	(4.7 to 21.0)	15	(8.2 to 25.0)
60	6	(2.3 to 15.0)	6.8	(2.6 to 17.0)	24	(12.0 to 44.0)	36	(21.0 to 56.0)
70	16	(6.4 to 38.0)	13	(5.2 to 32.0)	38	(21.0 to 60.0)	54	(32.0 to 78.0)
80	33	(14.0 to 66.0)	22	(8.9 to 50.0)	51	(30.0 to 76.0)	67	(42.0 to 90.0)

It was assumed that no competing risks took effect.
CRC, colorectal carcinoma; EC, endometrial carcinoma; MC, minor HNPCC cancer site.

with a person who did not have any previous event. A person who might have died from a cancer not included in the competing risks would cause a risk difference in the other direction.

Families selected on many cancer cases are likely not only to contain a high proportion of mutational carriers but also to overrepresent other genetic and environmental risk factors, as well as high penetrance mutations.[25] As a consequence this will wrongly attribute more risk to the mutation under study. Such an overestimated mutation risk would still be representative for mutation carriers from high risk families, to which we would like to give genetic counselling.[26] In our study, we assumed the population risk for persons without a mutation. Sensitivity analysis gave no evidence for excess familial risk.

A follow up study of 199 non-diseased mutation carriers found a risk for developing CRC under endoscopic surveillance of 10.5% within 10 years, corresponding to an annual incidence of 1.0% (95% confidence interval, 3.8 to 17.2).[27] Taking into account a considerable risk reduction resulting from screening, this would suggest high lifetime risks for CRC. In our study, an incidence above 1.0% was only found at age 50 and 60 for men and for an even narrower age range in women (table 4, fig 1). However, one has to consider the high degree of uncertainty within both studies.

The bell shape of the incidence function (fig 1) was caused by the data driven choice of the relative hazard function. The decline in CRC incidence for mutation carriers below the population incidence probably reflects the lack of data at higher ages.

Conclusions

The CRC risk estimates for mutation carriers in the current study are the lowest reported so far. These lower estimates might have an impact on counselling. However, we do not believe that there is a need to change screening practice, because the CRC risks are still high. Additional studies are needed to assess the impact of other genetic and environmental factors.

ACKNOWLEDGEMENTS

The electronic supplementary appendix can be found on the JMG web site, www.jmedgenet.com/supplemental

....

Authors' affiliations

F Quehenberger, H C van Houwelingen, Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, Netherlands
H F A Vasen, The Netherlands Foundation for the Detection of Hereditary Tumours, Leiden University Medical Centre

Competing interests: none declared

Correspondence to: Dr Hans F A Vasen, The Netherlands Foundation for the Detection of Hereditary Tumours, Leiden University Medical Centre, Rijnsburgerweg 10, ''Poortgebouw Zuid'', 2333 AA Leiden, Netherlands; nfdht@xs4all.nl

REFERENCES

1 **Lynch HT**, Smyrk T, Lynch J. An update of HNPCC (Lynch syndrome). *Cancer Genet Cytogenet* 1997;**93**:84–99.
2 **Vasen HF**, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* 1999;**116**:1453–6.
3 **Syngal S**, Fox EA, Eng C, Kolodner RD, Garber JE. Sensitivity and specificity of clinical criteria for hereditary non-polyposis colorectal cancer associated mutations in MSH2 and MLH1. *J Med Genet* 2000;**37**:641–5.
4 **Carayol J**, Khlat M, J JM, Bonaiti-Pellie C. Hereditary non-polyposis colorectal cancer: current risks of colorectal cancer largely overestimated. *J Med Genet* 2002;**39**:335–9.
5 **Salovaara R**, Loukola A, Kristo P, Kaariainen H, Ahtola H, Eskelinen M, Harkonen N, Julkunen R, Kangas E, Ojala S, Tulikoura J, Valkamo E, Jarvinen H, Mecklin JP, Aaltonen LA, de la Chapelle A. Population-based molecular detection of hereditary non-polyposis colorectal cancer (published erratum appears in *J Clin Oncol* 2000;18:3456). *J Clin Oncol* 2000;**18**:2193–200.
6 **Dunlop MG**, Farrington SM, Nicholl I, Aaltonen L, Petersen G, Porteous M, Carothers A. Population carrier frequency of hMSH2 and hMLH1 mutations. *Br J Cancer* 2000;**83**:1643–5.
7 **Gail MH**, Pee D, Benichou J, Carroll R. Designing studies to estimate the penetrance of an identified autosomal dominant mutation: cohort, case–control, and genotyped-proband designs. *Genet Epidemiol* 1999;**16**:15–39.
8 **Wacholder S**, Hartge P, Struewing JP, Pee D, McAdams M, Brody LM, Tucker M. The kin-cohort study for estimating penetrance. *Am J Epidemiol* 1998;**148**:623–30.
9 **Dunlop MG**, Farrington SM, Carothers AD, Wyllie AH, Sharp L, Burn J, Liu B, Kinzler KW, Vogelstein B. Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum Mol Genet* 1997;**6**:105–10.
10 **Shute NC**, Ewens WJ. A resolution of the ascertainment sampling problem. III. Pedigrees. *Am J Hum Genet* 1988;**43**:387–95.
11 **Hodge S**, Elston R. Lods, wrods, and mods: the interpretation of lod scores calculated under different models. *Genet Epidemiol* 1994;**11**:329–42.
12 **Vasen HF**, den Hartog Jager FC, Menko FH, Nagengast FM. Screening for hereditary non-polyposis colorectal cancer: a study of 22 kindreds in The Netherlands. *Am J Med* 1989;**86**:278–81.
13 **Vasen HFA**, Stormorken A, Menko FH, Nagengast FM, Kleibeuker JH, Griffioen G, Taal BG, Moller P, Wijnen JT. MSH2 mutation carriers are at higher risk of cancer than MLH1 mutation carriers: A study of hereditary non-polyposis colorectal cancer families. *J Clin Oncol* 2001;**19**:4074–80.
14 **Kalbeisch JD**, Prentice RL. *The statistical analysis of failure time data.* New York: John Wiley, 1980.
15 **Carayol J**, Bonaiti-Pellie C. Estimating penetrance from family data using a retrospective likelihood when ascertainment depends on genotype and age of onset. *Genet Epidemiol* 2002;**27**:109–17.
16 **Visser O**, Coebergh JWW, van Dijck JAAM, Siesling S, eds. *Incidence of cancer in the Netherlands 1998.* Utrecht: Vereniging van Integrale Kankercentra, 2002.
17 **Lange K**, Weeks D, Boehnke M. Programs for pedigree analysis – MENDEL, FISHER, and DGENE. *Genet Epidemiol* 1988;**5**:471–2.
18 **Froggatt NJ**, Brassett C, Koch DJ, Evans DGR, Hodgson SV, Ponder BAJ, Maher ER. Mutation screening of MSH2 and MLH1 mRNA in hereditary non-polyposis colon cancer syndrome. *J Med Genet* 1996;**33**:726–30.
19 **Vasen HFA**, Wijnen JT, Menko FH, Kleibeuker JH, Taal BG, Griffioen G, Nagengast FM, MeijersHeijboer EH, Bertario L, Varesco L, Bisgaard ML, Mohr J, Fodde R, Khan PM. Cancer risk in families with hereditary non-polyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology* 1996;**110**:1020–7.
20 **Aarnio M**, Mecklin JP, Aaltonen LA, NystromLahti M, Jarvinen HJ. Lifetime risk of different cancers in hereditary non-polyposis colorectal cancer (HNPCC) syndrome. *Int J Cancer* 1995;**64**:430–3.
21 **Aarnio M**, Sankila R, Pukkala E, Salovaara R, Aaltonen LA, de la Chapelle A, Peltomaki P, Mecklin JP, Jarvinen HJ. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer* 1999;**81**:214–18.
22 **Parc Y**, Boisson C, Thomas G, Olschwang S. Cancer risk in 348 French MSH2 or MLH1 gene carriers. *J Med Genet* 2003;**40**:208–13.
23 **Watson P**, Lynch HT. Cancer risk in mismatch repair gene mutation carriers. *Fam Cancer* 2001;**1**:57–60.
24 **Lin KM**, Shashidharan M, Ternent CA, Thorson AG, Blatchford GJ, Christensen MA, Lanspa SJ, Lemon SJ, Watson P, Lynch HT. Colorectal and extra-colonic cancer variations in MLH1/MSH2 hereditary non-polyposis colorectal cancer kindreds and the general population. *Dis Colon Rectum* 1998;**41**:428–33.
25 **Begg CB**. On the use of familial aggregation in population-based case probands for calculating penetrance. *J Natl Cancer Inst* 2002;**94**:1221–6.
26 **Burton PR**, Palmer LJ, Jacobs K, Keen KJ, Olson JM, Elston RC. Ascertainment adjustment: where does it take us? (erratum in Am J Hum Genet 2001;69:672). *Am J Hum Genet* 2000;**67**:1505–14.
27 **de Vos tot Nederveen Cappel WH**, Nagengast FM, Griffioen G, Menko FH, Taal BG, Kleibeuker JH, Vasen HF. Surveillance for hereditary non-polyposis colorectal cancer: a long-term study on 114 families. *Dis Colon Rectum* 2002;**45**:1588–94.