Germline variants in the mismatch repair genes: Detection and phenotype

Suerink, M.

Part II

Pheno type

An alternative approach to establishing unbiased colorectal cancer risk estimation in Lynch syndrome

Genetics in Medicine, 2019

Manon Suerink, Mar Rodríguez-Girondo, Heleen M. van der Klift, Chrystelle Colas, Laurence Brugieres, Noémie Lavoine, Marjolijn Jongmans, Gabriel Capellá Munar, D. Gareth Evans, Michael P. Farrell, Maurizio Genuardi, Yael Goldberg, Encarna Gomez-Garcia, Karl Heinimann,, Jessica I. Hoell, Stefan Aretz, Kory W. Jasperson, Inbal Kedar, Mitul B. Modi, Sergey Nikolaev, Theo A. M. van Os, Tim Ripperger, Daniel Rueda, Leigha Senter, Wenche Sjursen, Lone Sunde, Christina Therkildsen, Maria G. Tibiletti, Alison H. Trainer, Yvonne J. Vos, Anja Wagner, Ingrid Winship, Katharina Wimmer, Stefanie Y. Zimmermann, Hans F. Vasen, Christi J. van Asperen, Jeanine J. Houwing-Duistermaat, Sanne W. ten Broeke, and Maartje Nielsen

ABSTRACT

Purpose

Biallelic pathogenic variants in the mismatch repair (MMR) genes cause a recessive childhood cancer predisposition syndrome known as constitutional mismatch repair deficiency (CMMRD). Family members with a heterozygous MMR variant have Lynch syndrome. We aimed at estimating cancer risk in these heterozygous carriers as a novel approach to avoid complicated statistical methods to correct for ascertainment bias.

Methods

Cumulative colorectal cancer incidence was estimated in a cohort of *PMS2*- and *MSH6*-associated families, ascertained by the CMMRD phenotype of the index, by using mutation probabilities based on kinship coefficients as analytical weights in a proportional hazard regression on the cause-specific hazards. Confidence intervals (CIs) were obtained by bootstrapping at the family level.

Results

The estimated cumulative colorectal cancer risk at age 70 years for heterozygous *PMS2* variant carriers was 8.7% (95% CI 4.3–12.7%) for both sexes combined, and 9.9% (95% CI 4.9–15.3%) for men and 5.9% (95% CI 1.6–11.1%) for women separately. For heterozygous *MSH6* variant carriers these estimates are 11.8% (95% CI 4.5–22.7%) for both sexes combined, 10.0% (95% CI 1.83–24.5%) for men and 11.7% (95% CI 2.10–26.5%) for women.

Conclusion

Our findings are consistent with previous reports that used more complex statistical methods to correct for ascertainment bias. These results underline the need for MMR gene–specific surveillance protocols for Lynch syndrome.

INTRODUCTION

Lynch syndrome (MIM 120435) is an inherited autosomal dominant condition predisposing to the development of primarily colorectal and endometrial cancer. It is caused by pathogenic variants in the mismatch repair (MMR) genes *MLH1* (MIM *120436), *MSH2* (MIM *609309), *MSH6* (MIM *600678), and *PMS2* (MIM *600259). Estimation of Lynch syndrome–associated cancer risk is challenging because until recently, testing for Lynch syndrome was based on clinical or family history criteria such as the Amsterdam II criteria and the (revised) Bethesda guidelines.[1,2] Consequently the majority of known Lynch syndrome families were ascertained based on familial cancer history. In recent years there has been a shift toward universal screening of all colorectal and endometrial cancer patients for tumor hallmarks of Lynch syndrome.[3,4] These hallmarks include aberrant immunohistochemistry for the MMR proteins and the presence of microsatellite instability.[5,6] Furthermore, panel testing of cancer genes, including the MMR genes, is becoming standard practice and is also performed in families with a cancer history that does not necessarily include Lynch syndrome–associated cancers.[7] Families identified through universal screening or panel testing may show lower penetrance for Lynch syndrome–associated malignancies, and Hampel et al. were among the first to notice that Lynch syndrome cancer risks are not as high as previously estimated based on analyses of families ascertained using existing guidelines.[8] Appropriate surveillance measures for these newly identified families can only be established if risks can be estimated accurately.

Based on retrospective cohorts, current estimates of lifetime colorectal cancer risks for carriers of pathogenic variants in *MLH1* and *MSH2* are between 52% and 97%.[9] Colorectal cancer risk estimates are lower for carriers of a pathogenic variant in *MSH6* (22–36%) and lowest of all for *PMS2* (11–20%).[9-12] A recent study of a prospective cohort of pathogenic MMR variant carriers undergoing surveillance reported even lower risks, with colorectal cancer risks of 12% for *MSH6* and 0% for *PMS2*, respectively.[13] As in the general population, men with Lynch syndrome appear to have a higher colorectal cancer risk than women.[14] In most studies, statistical approaches such as modified segregation analysis, exclusion of index cases, and genotype-restricted likelihood estimates have been used to correct for ascertainment bias, but these methods are complex and rely on specific assumptions, and it is difficult to prove that they do not lead to either under- or overestimation of true risk.[14] Indeed, Vos et al. showed that a substantial proportion of the variation found in cancer risk estimation in selected hereditary breast cancer families, who show similar ascertainment patterns to Lynch syndrome families, can be explained by the different ascertainment correction method used.[15]

An alternative approach that minimizes the need for ascertainment bias correction is the selection of families in which the index patient has constitutional mismatch repair deficiency (CMMRD). This childhood cancer predisposition syndrome is caused by biallelic pathogenic variants in one of the MMR genes, most commonly in *PMS2*. The syndrome is characterized by the development of a broad spectrum of cancers, including hematological, central nervous system, and gastrointestinal neoplasia at a very young age. CMMRD patients may also show signs suggestive of neurofibromatosis type 1, most commonly café au lait macules.[16] The CMMRD phenotype is so striking that the diagnosis is often suspected regardless of family history and in one report only 6 of 23 CMMRD patients (26%) had a family history of Lynch syndrome–associated cancers.[17] Identification of a child with CMMRD means that both parents are likely to be heterozygous for a pathogenic MMR variant and are at risk for Lynch syndrome–associated malignancies; other family members may similarly be at risk. Because these families were identified due to the CMMRD phenotype rather than family history, they likely represent a near random sample of Lynch syndrome families.

Pathogenic variants in *PMS2* were once considered rare and were thought to account for less than 5% of all Lynch syndrome cases.[18,19] Nevertheless, germline pathogenic variants in PMS2 were found in a small yet significant proportion (at least 0.57%) of universally screened colorectal cancer cases,[20] and recent insights suggest that the carrier frequency for pathogenic variants in *PMS2* and *MSH6* in the general population is actually much higher than for *MLH1* and *MSH2*.[21] The majority of CMMRD patients carry variants in *PMS2*, followed by *MSH6*, while *MLH1* and *MSH2* variants are rarely associated with CMMRD.[16] One explanation for this phenomenon is that biallelic pathogenic variants in *MLH1* and *MSH2* may be embryonically lethal.[22,23] However, a higher carrier frequency for variants in *PMS2* and *MSH6* may also (partly) explain differences in the frequency of pathogenic variants in the MMR genes among patients with CMMRD.

Here we report cumulative cancer risks in family members of CMMRD patients with variants in the *PMS2* or *MSH6* genes. This study will not only help in the counseling of family members of CMMRD patients, but also represents a novel approach to determining cancer risk in Lynch syndrome.

MATERIALS AND METHODS

Data collection

Families were collected through international collaborations with clinical genetics departments and consortia and by following up CMMRD families described in literature. Corresponding authors were contacted to collect (more) family data. Family structure was recorded and information was collected on each family member regarding gender, variant status, cancer status and age at cancer diagnosis, and last contact or death. A diagnosis of CMMRD was considered confirmed if pathogenic variants were identified or if strong indicators of CMMRD were identified (i.e., phenotype and inheritance pattern plus aberrant immunohistochemistry and/ or microsatellite instability in non-neoplastic tissue and/or abnormal functional tests).[24]

As classified in the InSiGHT database (http://www.insight-database.org/classifications/), 31 unique class 4/5 pathogenic variants in *PMS2* and 19 class 4/5 pathogenic variants in *MSH6* were found in our cohort.[25] Another 30 variants in *PMS2* and 8 variants in *MSH6* have not been officially classified to date, but were deemed either class 4 or 5 (i.e., [likely] pathogenic) by an expert in the field (H.M.v. d.K.) according to InSiGHT variant classification criteria. Twenty variants of uncertain significance (VUS), distributed over 18 families, were identified and included in the analyses (Tables S1–S4). Seven of the VUS were identified in trans with a (likely) pathogenic variant. Since the patients carrying these VUS displayed a CMMRD phenotype this argues in favor of a functional impact of the variants on protein function. Furthermore, six of the VUS were identified in previously published CMMRD patients (Tables S3 and S4) and as such these variants were considered the most probable cause of the phenotype in these patients. The remaining seven variants were all identified in patients with a CMMRD phenotype and were considered a probable cause of the phenotype by the reporting laboratory and clinicians.

Statistical analysis

Eligible first- and second-degree family members for the risk analysis were defined based on complete data describing gender, age at cancer diagnosis, last contact or death, and status as a (possible) carrier of the *PMS2*/*MSH6* variant. Proven and obligate carriers as well as untested family members were included, whereas noncarriers, as confirmed by DNA analysis, were excluded. Known CMMRD patients were excluded from the analysis, as were (deceased) siblings of a CMMRD patient when they had a cancer within the CMMRD spectrum. In consanguineous families, family members with an unknown variant status, but a cancer diagnosis within the CMMRD cancer spectrum

at a young age (i.e., <25 years of age) were considered to be homozygous carriers and were thus excluded from the risk analysis. The total number of colorectal and endometrial cancers is described for the total cohort as well as for the part of the cohort included in the risk analysis. To avoid a reporting bias due to distant relatives (distant family members may be more likely to be included in the pedigree if they were affected, while unaffected distant family members may go unreported), only first- and second-degree relatives of the index patients were included in the risk analyses. This approach was supported by both visual inspection of the pedigrees and by an otherwise unexplained increase in colorectal cancer frequency among more distant family members (data not shown, available upon request).

Colorectal cancer risk is reported as cumulative incidence at age 70, accounting for death and other cancer diagnoses as competing risks.[26] Age at removal of a colon polyp was included as a censoring event because the likelihood of developing colorectal cancer is probably reduced after this preventive measure. Likewise, family members were censored at the development of any type of cancer, excluding basal cell carcinoma, because treatment of a cancer (e.g., by radiotherapy or chemotherapy) might influence future cancer risk.

To avoid testing bias, which may arise when the decision to undergo genetic testing is related to cancer status, we included untested family members in our study, weighted according to their genetic distance to confirmed carriers. Specifically, variant probabilities based on kinship coefficients were used as analytical weights in a Cox proportional hazard regression to model the hazard of developing colorectal cancer in the presence of competing events (death and other cancer diagnosis), and including sex as a covariate (for details see "Statistical Methods" in the Supplemental Data). For example, first-degree relatives of a confirmed carrier who were not tested were given a weight of 0.50, whereas second-degree relatives had a weight of 0.25. Confidence intervals (CIs) were obtained by bootstrapping at family level (1000 repetitions).

Medical ethical approval for this study was obtained through the ethics committee of Leiden University Medical Centre (reference number P14.090). Informed consent was not required because all data was collected anonymously.

RESULTS

After exclusion of the CMMRD cases, the *PMS2* cohort included 1809 family members from 77 families and the *MSH6* cohort consisted of 561 family members from 26 families.

Age at colorectal and endometrial cancer diagnosis

Sixty patients from 31 families were diagnosed with colorectal cancer in the total *PMS2* cohort, and 16 women from 14 families were diagnosed with endometrial cancer after excluding the CMMRD cases. Age of colorectal cancer diagnosis within this cohort ranged from 36 to 80 years, with a median age of 60 years. Age at diagnosis was unknown for 17 colorectal cancer cases (Table 1). For the 16 endometrial cancer cases, the age at diagnosis ranged from 40 to 85, with a median of 61 years. Age was missing for only one of these cases.

Seventeen patients from 12 families were diagnosed with colorectal cancer in the total *MSH6* cohort after exclusion of CMMRD cases. Age of colorectal cancer diagnosis in this cohort ranged from 42 to 58 years, with a median of 48 years (Table 1). There were five cases of endometrial cancer distributed over four families, with a median age at diagnosis of 54 years and an age range of 47 to 59 years.

Table 1. Cohort description, CMMRD patients excluded. CRC = colorectal cancer, EC = endometrial cancer

gene			total cohort	in risk analysis
PMS2	number of family members		1809	549
	gender	male female unknown	858 (47.4%) 728 (40.2%) 223 (12.3%)	299 (51.7%) 283 (48.3%) -
	carrier status	carrier unknown	369 1440	212 337
	age (years)	median (range) missing (n)	43.0 (0-94) 1235	49.0 (0-93) -
	CRC	n	60	21
	age at CRC diagnosis (years)	median (range) missing (n)	60.0 (36-80) 17	60.0 (36-80) -
	competing events (right censoring)			
	EC	n	16	6
	age at EC diagnosis (years)	median (range) missing (n)	61.0 (40-85) 1	61.5 (50-80) -
	other cancer or polypectomy/ hysterectomy	n	85	6
	age at other cancer diagnosis or removal of first polyp or uterus (years)	median (range) missing (n)	55.0 (5-85) 11	54 (5-84) -
	death	n	112	44
	age at death (years)	median (range) missing (n)	69.0 (0-94) 55	68.5 (0-93) -
MSH6	number of family members		561	148
	gender	male female unknown	299 (53.3%) 252 (44.9%) 10 (1.8%)	76 (51.4%) 72 (48.6%) -
	carrier status	carrier unknown	146 415	69 79
	age (years)	median (range) missing (n)	43.0 (3-86) 336	45.0 (1-85) -
	CRC	n	17	8
	age at CRC diagnosis (years)	median (range) missing (n)	48.0 (42-58) 4	47.5 (42-58) -
	competing events (right censoring)			
	EC	n	5	0
	age at EC diagnosis (years)	median (range) missing (n)	54.0 (47-59) 0	Not applicable
	other cancer or polypectomy/ hysterectomy	n	40	25
	age at other cancer diagnosis or removal of first polyp (years)	median (range) missing (n)	52.0 (7-78) 3	57.0 (23-78) -
	death	n	37	11
	age at death (years)	median (range) missing (n)	38.5 (1-81) 1	25.0 (1-73) -

Other cancers

While a range of other cancer types were reported in both the *PMS2* and *MSH6* cohort, low numbers did not allow risk analyses to be performed. The most commonly reported cancers were breast cancer, lung cancer, leukemia, and prostate cancer (Table 1 and Table S5).

Colorectal cancer risk

For individuals with CMMRD and variants in *PMS2*, 549 family members from 64 families were eligible for risk analysis; of these, 212 were confirmed or obligate carriers and the rest potential carriers. The estimated cumulative colorectal cancer risk at age 70 for heterozygous *PMS2* variant carriers was 8.7% (95% CI 4.3–12.7%, Fig. 1) for both sexes combined, and was 9.9% (95% CI 4.9–15.3%) for men and 5.9% (95% CI 1.6–11.1%) for women. Endometrial cancer risk could not be estimated due to the low number of events (n = 8).

For *MSH6*, 148 family members from 24 families were eligible for risk analysis; of these 69 were confirmed or obligate carriers and the rest potential carriers. The cumulative colorectal cancer risk at age 70 for heterozygous *MSH6* gene variant carriers was 11.8% (95% CI 4.5–22.7%, Fig. 2) for both sexes, and 10.0% (95% CI 1.8–24.5%) and 11.7% (95% CI 2.1–26.5%) for men and women, respectively. There were no cases of endometrial cancer that could be included in the risk analysis.

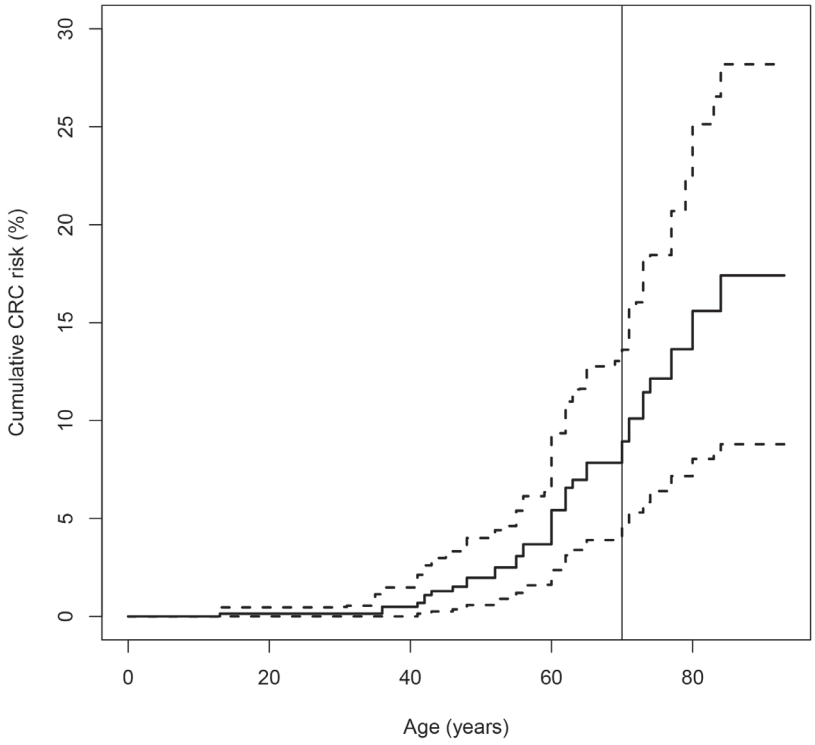

Figure 1 Cumulative colorectal cancer risk for carriers of a pathogenic PMS2 variant, men and women together, with 95% confidence intervals shown as dashed lines. CRC = colorectal cancer.

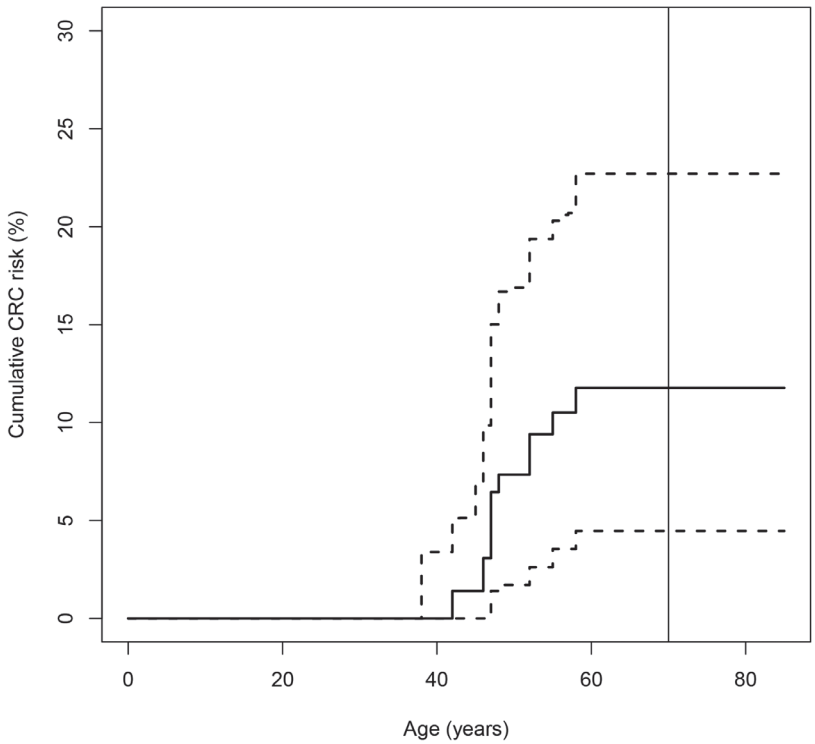

Figure 2 Cumulative colorectal cancer risk for carriers of a pathogenic MSH6 variant, men and women together, with 95% confidence intervals shown as dashed lines. CRC = colorectal cancer.

DISCUSSION

Using a new approach to establishing cancer risks in Lynch syndrome, we can confirm the low *PMS2*- and *MSH6*-associated colorectal cancer risks reported in previous studies that used ascertainment bias correction methods[10-12,14] or prospective data.[13,27] The main strengths of our approach were the reduction in clinical ascertainment bias by analyzing family members of CMMRD patients and the use of a competing risk analysis approach to avoid bias due to informative right censoring. Our results further indicate that gene-specific surveillance guidelines are needed to avoid subjecting carriers at low cancer risk to the invasive processes of surveillance, in some cases from an unnecessarily young age. The earliest age of colorectal cancer diagnosis was 36 and 42 years for *PMS2* and *MSH6*, respectively, well above the age (20–25 years) at which surveillance is usually started for individuals with Lynch syndrome.[28] This suggests that, in heterozygous carriers of *PMS2* or *MSH6* variants from families that do not meet clinical selection criteria for Lynch syndrome, surveillance could be started at a later age, e.g., at 35–40 years. Although current lifetime risk estimates are only slightly (2–3 times) elevated above the population risk of ~4%,[29] there are indications (e.g., from the median age at diagnosis) that risk is elevated at younger ages, and a faster progression from precursor lesion to carcinoma cannot be excluded. Therefore, we do not recommend that surveillance be omitted based on the current data. Furthermore, large variation in penetrance has been observed in clinically ascertained families, indicating that other risk factors may influence risk. Together these considerations suggest that our risk estimates remain useful when counseling families who were not ascertained based on criteria such as the Amsterdam II criteria and the (revised) Bethesda guidelines, e.g., families with a CMMRD proband or with a pathogenic MMR variant identified as an incidental finding through exome sequencing. However, they should be used with caution in more severely affected families, for example when a family history fulfills the Amsterdam criteria.[2]

Unfortunately, both cohorts were too small to provide risk estimations for endometrial cancer. It is striking that there were only some cases of endometrial cancer in the total *MSH6* cohort and none that could be included in the risk analysis, while the risk of endometrial cancer in *MSH6* has been reported to be high.[27] This may be partly due to the relatively low median age of 45 years (Table 1) of the cohort, while the youngest age at diagnosis of endometrial cancer was 47 years (Table 1).

There are some limitations to the current study. Firstly, genotype–phenotype correlations in Lynch syndrome and CMMRD have been proposed (although thus far no conclusive evidence has been yielded and some studies even show contradictory results).[30-35] If

correlations exist, variants with a milder phenotype might be overrepresented in a CMMRD cohort. For *PMS2*, age at cancer diagnosis and risk estimates were within the range of previous retrospective studies that corrected for ascertainment bias, indicating that we have not selected a cohort of (solely) low-risk PMS2 alleles.[10-12] Cancer risk estimates and age at cancer diagnosis for *MSH6* are similar to a study by Bonadona et al.,[36] but risk estimates are slightly lower than those reported by Baglietto et al.[37]

A possible mechanism for a genotype–phenotype correlation could be nonsense-mediated messenger RNA (mRNA) decay. Nonsense-mediated decay (NMD) detects mRNAs with premature termination codons and initiates their degradation, preventing potential dominant negative effects from truncated proteins.[38] Some variants, e.g., missense variants, are likely to escape NMD. To assess a possible role for NMD, we performed a stratified risk analysis that divided family members into groups based on whether their risk variant is expected to result in NMD, as described previously (Suerink et al.[30]). Family members were excluded from this analysis when no reliable prediction of NMD was available for the variant or if it was not known which variant segregated in which half of the family (maternal or paternal). This analysis produced no clear genotype–phenotype correlations and for both genes cases of colorectal cancer were seen in the NMD group as well as in the group with predicted retention of RNA expression. However, it should be noted that wide confidence intervals excluded detection of small differences (data available upon request). Whether risk stratification is possible based on genotype will require further study.

It could also be argued that a bias toward a milder phenotype is inherent to our cohort because those who die of cancer at a young age cannot have children with CMMRD. However, because both the parents and more distant relatives were included in the current analyses, it seems unlikely that this possible bias could have a major impact, particularly because the youngest age at colorectal cancer diagnosis within the total cohort was 36 years. Another potential problem was testing bias, which arises because family members with cancer are more inclined to undergo genetic testing. We therefore used variant probabilities based on genetic distance to confirmed carriers as analytical weights in our statistical analysis, which also enabled inclusion of untested family members. By including obligate carriers in the analysis there is a risk of misidentifying someone as a possible carrier because the CMMRD patient may have had a de novo variant. However, de novo variants are rarely reported in Lynch syndrome (2.3% in a cohort described by Win et al.[39]) and a large proportion (55% and 50% for *PMS2* and *MSH6*, respectively) of CMMRD index patients were homozygous for one variant and/or were from consanguineous families. Moreover, a major testing bias was not expected due to

a low overall cancer risk and because a relatively large proportion of confirmed carriers were obligate carriers (45/212 [21%] for the *PMS2* cohort and 21/69 [30%] for the *MSH6* cohort) whose testing status is by definition uninfluenced by their phenotype. It is worth mentioning that while our approach avoids clinical ascertainment bias, the selection strategy results in a relatively young cohort, which implies large uncertainty in the incidence estimation at older ages, as reflected by the broad confidence intervals in Figs. 1 and 2.

A final limitation of our study that could impact the reliability of data is the fact that most cancer diagnoses in this cohort were based on the proband's knowledge of family history rather than on medical records. Reassuringly, a 2011 study showed that the accuracy of reported colorectal cancer for first-degree family members was over 90%.[40] Because we included only first and second-degree family members, with family history reported by the parents in most cases, we expect a comparable accuracy rate in our risk analysis.

To complement and confirm the data presented here, we suggest a similar risk analysis should be performed in *PMS2* and *MSH6* families detected through universal screening of colorectal cancers for mismatch repair deficiency. These families will also be less affected with ascertainment bias.

In summary, we used an alternative approach to establish colorectal cancer risk in Lynch syndrome patients with *PMS2* and *MSH6* variants in CMMRD families. We confirmed this relatively low cancer risk relative to earlier, biased estimates of risk. These results underline the need for gene-specific surveillance protocols for *PMS2*- and *MSH6*-related Lynch syndrome families. Further investigations will be required to estimate the cancer risk for other Lynch syndrome–associated malignancies for *PMS2* and *MSH6*, as well as estimating unbiased cancer risks estimates for carriers of pathogenic variants in *MLH1* and *MSH2*.

ACKNOWLEDGEMENTS

We acknowledge Susan E. Andrew (Department of Medical Genetics, University of Alberta, Edmonton, Canada) and Kate Green (Division of Evolution and Genomic Medicine, Manchester Academic Health Science Centre [MAHSC], University of Manchester, St Mary's Hospital, Manchester, UK) for providing data, and the Care for CMMRD (C4CMMRD) Consortium for providing data and a platform to discuss this study. The authors thank Medactie.com for help with editing of this paper. This work was supported by a grant from the Dutch Cancer Society (KWF UL 2012–5155).

REFERENCES

1. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *Journal of the National Cancer Institute*. 2004;96(4):261-268.
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(6):1453-1456.
3. Leenen CH, Goverde A, de Bekker-Grob EW, Wagner A, van Lier MG, Spaander MC, Bruno MJ, Tops CM, van den Ouweland AM, Dubbink HJ, Kuipers EJ, Dinjens WN, van Leerdam ME, Steyerberg EW. Cost-effectiveness of routine screening for Lynch syndrome in colorectal cancer patients up to 70 years of age. *Genetics in Medicine*. 2016;18(10):966-973.
4. Dillon JL, Gonzalez JL, DeMars L, Bloch KJ, Tafe LJ. Universal screening for Lynch syndrome in endometrial cancers: frequency of germline mutations and identification of patients with Lynch-like syndrome. *Human Pathology*. 2017;70:121-128.
5. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Research*. 1998;58(22):5248-5257.
6. Richman S. Deficient mismatch repair: Read all about it (Review). *International Journal of Oncology*. 2015;47(4):1189-1202.
7. Desmond A, Kurian AW, Gabree M, Mills MA, Anderson MJ, Kobayashi Y, Horick N, Yang S, Shannon KM, Tung N, Ford JM, Lincoln SE, Ellisen LW. Clinical Actionability of Multigene Panel Testing for Hereditary Breast and Ovarian Cancer Risk Assessment. *JAMA Oncol*. 2015;1(7):943-951.
8. Hampel H, Stephens JA, Pukkala E, Sankila R, Aaltonen LA, Mecklin JP, de la Chapelle A. Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology*. 2005;129(2):415-421.
9. Hampel H, de la Chapelle A. The search for unaffected individuals with Lynch syndrome: do the ends justify the means? *Cancer Prevention Research (Philadelphia, Pa)*. 2011;4(1):1-5.
10. Senter L, Clendenning M, Sotamaa K, Hampel H, Green J, Potter JD, Lindblom A, Lagerstedt K, Thibodeau SN, Lindor NM, Young J, Winship I, Dowty JG, White DM, Hopper JL, Baglietto L, Jenkins MA, de la Chapelle A. The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. *Gastroenterology*. 2008;135(2):419-428.
11. ten Broeke SW, Brohet RM, Tops CM, van der Klift HM, Velthuizen ME, Bernstein I, Capella Munar G, Gomez Garcia E, Hoogerbrugge N, Letteboer TG, Menko FH, Lindblom A, Mensenkamp AR, Moller P, van Os TA, Rahner N, Redeker BJ, Sijmons RH, Spruijt L, Suerink M, Vos YJ, Wagner A, Hes FJ, Vasen HF, Nielsen M, Wijnen JT. Lynch syndrome caused by germline PMS2 mutations: delineating the cancer risk. *Journal of Clinical Oncology*. 2015;33(4):319-325.
12. Ten Broeke SW, van der Klift HM, Tops CMJ, Aretz S, Bernstein I, Buchanan DD, de la Chapelle A, Capella G, Clendenning M, Engel C, Gallinger S, Gomez Garcia E, Figueiredo JC, Haile R, Hampel HL, Hopper JL, Hoogerbrugge N, von Knebel Doeberitz M, Le Marchand L, Letteboer TGW, Jenkins MA, Lindblom A, Lindor NM, Mensenkamp AR, Moller P, Newcomb PA, van Os TAM, Pearlman R, Pineda M, Rahner N, Redeker EJW, Olderode-Berends MJW, Rosty C, Schackert HK, Scott R, Senter L, Spruijt L, Steinke-Lange V, Suerink M, Thibodeau S, Vos YJ, Wagner A, Winship I, Hes FJ, Vasen HFA, Wijnen JT, Nielsen M, Win AK. Cancer Risks for PMS2-Associated Lynch Syndrome. *Journal of Clinical Oncology*. 2018;36(29):2961-2968.

13. Moller P, Seppala T, Bernstein I, Holinski-Feder E, Sala P, Evans DG, Lindblom A, Macrae F, Blanco I, Sijmons R, Jeffries J, Vasen H, Burn J, Nakken S, Hovig E, Rodland EA, Tharmaratnam K, de Vos Tot Nederveen Cappel WH, Hill J, Wijnen J, Green K, Lalloo F, Sunde L, Mints M, Bertario L, Pineda M, Navarro M, Morak M, Renkonen-Sinisalo L, Frayling IM, Plazzer JP, Pylvanainen K, Sampson JR, Capella G, Mecklin JP, Moslein G, Mallorca G. Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database. *Gut.* 2017;66(3):464-472.
14. Barrow E, Hill J, Evans DG. Cancer risk in Lynch Syndrome. *Familial Cancer.* 2013;12(2):229-240.
15. Vos JR, Hsu L, Brohet RM, Mourits MJ, de Vries J, Malone KE, Oosterwijk JC, de Bock GH. Bias Correction Methods Explain Much of the Variation Seen in Breast Cancer Risks of BRCA1/2 Mutation Carriers. *Journal of Clinical Oncology.* 2015;33(23):2553-2562.
16. Wimmer K, Kratz CP, Vasen HF, Caron O, Colas C, Entz-Werle N, Gerdes AM, Goldberg Y, Ilencikova D, Muleris M, Duval A, Lavoine N, Ruiz-Ponte C, Slavc I, Burkhardt B, Brugieres L, CMMRD EU-CCf. Diagnostic criteria for constitutional mismatch repair deficiency syndrome: suggestions of the European consortium 'care for CMMRD' (C4CMMRD). *Journal of Medical Genetics.* 2014;51(6):355-365.
17. Lavoine N, Colas C, Muleris M, Bodo S, Duval A, Entz-Werle N, Coulet F, Cabaret O, Andreiuolo F, Charpy C, Sebille G, Wang Q, Lejeune S, Buisine MP, Leroux D, Couillault G, Leverger G, Fricker JP, Guimbaud R, Mathieu-Dramard M, Jedraszak G, Cohen-Hagenauer O, Guerrini-Rousseau L, Bourdeaut F, Grill J, Caron O, Baert-Dusermont S, Tinat J, Bougeard G, Frebourg T, Brugieres L. Constitutional mismatch repair deficiency syndrome: clinical description in a French cohort. *Journal of Medical Genetics.* 2015;52(11):770-778.
18. Gill S, Lindor NM, Burgart LJ, Smalley R, Leontovich O, French AJ, Goldberg RM, Sargent DJ, Jass JR, Hopper JL, Jenkins MA, Young J, Barker MA, Walsh MD, Ruszkiewicz AR, Thibodeau SN. Isolated loss of PMS2 expression in colorectal cancers: frequency, patient age, and familial aggregation. *Clinical Cancer Research.* 2005;11(18):6466-6471.
19. Peltomaki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Human Molecular Genetics.* 2001;10(7):735-740.
20. Truninger K, Menigatti M, Luz J, Russell A, Haider R, Gebbers JO, Bannwart F, Yurtsever H, Neuweiler J, Riehle HM, Cattaruzza MS, Heinimann K, Schar P, Jiricny J, Marra G. Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer. *Gastroenterology.* 2005;128(5):1160-1171.
21. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, Buchanan DD, Clendenning M, Rosty C, Ahnen DJ, Thibodeau SN, Casey G, Gallinger S, Le Marchand L, Haile RW, Potter JD, Zheng Y, Lindor NM, Newcomb PA, Hopper JL, MacInnis RJ. Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. *Cancer Epidemiology, Biomarkers and Prevention.* 2017;26(3):404-412.
22. Bakry D, Aronson M, Durno C, Rimawi H, Farah R, Alharbi QK, Alharbi M, Shamvil A, Ben-Shachar S, Mistry M, Constantini S, Dvir R, Qaddoumi I, Gallinger S, Lerner-Ellis J, Pollett A, Stephens D, Kelies S, Chao E, Malkin D, Bouffet E, Hawkins C, Tabori U. Genetic and clinical determinants of constitutional mismatch repair deficiency syndrome: report from the constitutional mismatch repair deficiency consortium. *European Journal of Cancer.* 2014;50(5):987-996.
23. Durno C, Boland CR, Cohen S, Dominitz JA, Giardiello FM, Johnson DA, Kaltenbach T, Levin TR, Lieberman D, Robertson DJ, Rex DK. Recommendations on Surveillance and Management of Biallelic Mismatch Repair Deficiency (BMMRD) Syndrome: A Consensus Statement by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology.* 2017;152(6):1605-1614.
24. Bodo S, Colas C, Buhard O, Collura A, Tinat J, Lavoine N, Guilloux A, Chalastanis A, Lafitte P, Coulet F, Buisine MP, Ilencikova D, Ruiz-Ponte C, Kinzel M, Grandjouan S, Brems H, Lejeune S, Blanche H, Wang Q, Caron O, Cabaret O, Svrcek M, Vidaud D, Parfait B, Verloes A, Knappe

UJ, Soubrier F, Mortemousque I, Leis A, Auclair-Perrossier J, Frebourg T, Flejou JF, Entz-Werle N, Leclerc J, Malka D, Cohen-Haguenauer O, Goldberg Y, Gerdes AM, Fedhila F, Mathieu-Dramard M, Hamelin R, Wafaa B, Gauthier-Villars M, Bourdeaut F, Sheridan E, Vasen H, Brugieres L, Wimmer K, Muleris M, Duval A, European Consortium "Care for C. Diagnosis of Constitutional Mismatch Repair-Deficiency Syndrome Based on Microsatellite Instability and Lymphocyte Tolerance to Methylating Agents. *Gastroenterology.* 2015;149(4):1017-1029 e1013.

25. Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, Bapat B, Bernstein I, Capella G, den Dunnen JT, du Sart D, Fabre A, Farrell MP, Farrington SM, Frayling IM, Frebourg T, Goldgar DE, Heinen CD, Holinski-Feder E, Kohonen-Corish M, Robinson KL, Leung SY, Martins A, Moller P, Morak M, Nystrom M, Peltomaki P, Pineda M, Qi M, Ramesar R, Rasmussen LJ, Royer-Pokora B, Scott RJ, Sijmons R, Tavtigian SV, Tops CM, Weber T, Wijnen J, Woods MO, Macrae F, Genuardi M. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nature Genetics.* 2014;46(2):107-115.
26. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine.* 2007;26(11):2389-2430.
27. Moller P, Seppala TT, Bernstein I, Holinski-Feder E, Sala P, Gareth Evans D, Lindblom A, Macrae F, Blanco I, Sijmons RH, Jeffries J, Vasen HFA, Burn J, Nakken S, Hovig E, Rodland EA, Tharmaratnam K, de Vos Tot Nederveen Cappel WH, Hill J, Wijnen JT, Jenkins MA, Green K, Lalloo F, Sunde L, Mints M, Bertario L, Pineda M, Navarro M, Morak M, Renkonen-Sinisalo L, Valentin MD, Frayling IM, Plazzer JP, Pylvanainen K, Genuardi M, Mecklin JP, Moeslein G, Sampson JR, Capella G, Mallorca G. Cancer risk and survival in path_MMR carriers by gene and gender up to 75 years of age: a report from the Prospective Lynch Syndrome Database. *Gut.* 2018;67(7):1306-1316.
28. de Vos tot Nederveen Cappel WH, Jarvinen HJ, Lynch PM, Engel C, Mecklin JP, Vasen HF. Colorectal surveillance in Lynch syndrome families. *Familial Cancer.* 2013;12(2):261-265.
29. Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA. SEER Cancer Statistics Review, 1975-2015. https://seer.cancer.gov/csr/1975_2015/, based on November 2017 SEER data submission, posted to the SEER web site, April 2018. Published 2018. Accessed2018.
30. Suerink M, van der Klift HM, Ten Broeke SW, Dekkers OM, Bernstein I, Capella Munar G, Gomez Garcia E, Hoogerbrugge N, Letteboer TG, Menko FH, Lindblom A, Mensenkamp A, Moller P, van Os TA, Rahner N, Redeker BJ, Olderode-Berends MJ, Spruijt L, Vos YJ, Wagner A, Morreau H, Hes FJ, Vasen HF, Tops CM, Wijnen JT, Nielsen M. The effect of genotypes and parent of origin on cancer risk and age of cancer development in PMS2 mutation carriers. *Genetics in Medicine.* 2016;18(4):405-409.
31. Geary J, Sasieni P, Houlston R, Izatt L, Eeles R, Payne SJ, Fisher S, Hodgson SV. Gene-related cancer spectrum in families with hereditary non-polyposis colorectal cancer (HNPCC). *Familial Cancer.* 2008;7(2):163-172.
32. Peltomaki P, Gao X, Mecklin JP. Genotype and phenotype in hereditary nonpolyposis colon cancer: a study of families with different vs. shared predisposing mutations. *Familial Cancer.* 2001;1(1):9-15.
33. Perez-Cabornero L, Infante M, Velasco E, Lastra E, Miner C, Duran M. Genotype-phenotype correlation in MMR mutation-positive families with Lynch syndrome. *International Journal of Colorectal Disease.* 2013;28(9):1195-1201.
34. Ryan NAJ, Morris J, Green K, Lalloo F, Woodward ER, Hill J, Crosbie EJ, Evans DG. Association of Mismatch Repair Mutation With Age at Cancer Onset in Lynch Syndrome: Implications for Stratified Surveillance Strategies. *JAMA Oncol.* 2017;3(12):1702-1706.
35. Li L, Hamel N, Baker K, McGuffin MJ, Couillard M, Gologan A, Marcus VA, Chodirker B, Chudley A, Stefanovici C, Durandy A, Hegele RA, Feng BJ, Goldgar DE, Zhu J, De Rosa M, Gruber SB,

Wimmer K, Young B, Chong G, Tischkowitz MD, Foulkes WD. A homozygous PMS2 founder mutation with an attenuated constitutional mismatch repair deficiency phenββotype. *Journal of Medical Genetics*. 2015;52(5):348-352.
36. Bonadona V, Bonaiti B, Olschwang S, Grandjouan S, Huiart L, Longy M, Guimbaud R, Buecher B, Bignon YJ, Caron O, Colas C, Nogues C, Lejeune-Dumoulin S, Olivier-Faivre L, Polycarpe-Osaer F, Nguyen TD, Desseigne F, Saurin JC, Berthet P, Leroux D, Duffour J, Manouvrier S, Frebourg T, Sobol H, Lasset C, Bonaiti-Pellie C, French Cancer Genetics N. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. *JAMA*. 2011;305(22):2304-2310.
37. Baglietto L, Lindor NM, Dowty JG, White DM, Wagner A, Gomez Garcia EB, Vriends AH, Dutch Lynch Syndrome Study G, Cartwright NR, Barnetson RA, Farrington SM, Tenesa A, Hampel H, Buchanan D, Arnold S, Young J, Walsh MD, Jass J, Macrae F, Antill Y, Winship IM, Giles GG, Goldblatt J, Parry S, Suthers G, Leggett B, Butz M, Aronson M, Poynter JN, Baron JA, Le Marchand L, Haile R, Gallinger S, Hopper JL, Potter J, de la Chapelle A, Vasen HF, Dunlop MG, Thibodeau SN, Jenkins MA. Risks of Lynch syndrome cancers for MSH6 mutation carriers. *Journal of the National Cancer Institute*. 2010;102(3):193-201.
38. Popp MW, Maquat LE. Nonsense-mediated mRNA Decay and Cancer. *Current Opinion in Genetics and Development*. 2018;48:44-50.
39. Win AK, Jenkins MA, Buchanan DD, Clendenning M, Young JP, Giles GG, Goldblatt J, Leggett BA, Hopper JL, Thibodeau SN, Lindor NM. Determining the frequency of de novo germline mutations in DNA mismatch repair genes. *Journal of Medical Genetics*. 2011;48(8):530-534.
40. Edwards E, Lucassen A. The impact of cancer pathology confirmation on clinical management of a family history of cancer. *Familial Cancer*. 2011;10(2):373-380.

Supplemental table 1. *PMS2* variants

PMS2 variant[a]	Change at RNA and/ or protein level[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
c.219T>A	p.(Cys73*)	nonsense	[5]	1 (0)	
c.400C>T	p.(Arg134*)	nonsense	5	1(1)	1 (1)
c.823C>T	r.823c>u, p.(Gln275*)	nonsense	[5]		1 (1)
c.862C>T	p.(Gln288*)	nonsense	[5]	1 (1)	
c.943C>T	p.(Arg315*)	nonsense	5	1 (1)	
c.949C>T	p.(Gln317*)	nonsense	5	1 (1)	
c.1840A>T	p.(Lys614*)	nonsense	5	1 (1)	
c.1882C>T	r.1882c>u, p.(Arg628*)	nonsense	5		1 (1)
c.1927C>T	p.(Gln643*)	nonsense	5		1 (1)
c.2192T>G	p.(Leu731*)	nonsense	[5]	1 (1)	
c.2404C>T	p.(Arg802*)	nonsense	5	2 (1)	
c.219_220dup	r.219_220dup, p.(Gly74Valfs*3)	frameshift	5	1 (1)	
c.247_250dup	r.247_250dup, p.(Thr84Ilefs*9)	frameshift	[5]		1 (1)
c.325dup	r.[325dup, 301_353del, 251_353del], p.([Glu109Glyfs*30, ?, ?])	frameshift	[5]		1 (1)
c.686_687del	p.(Ser229Cysfs*19)	frameshift	[5]	1 (1)	
c.736_741delinsTGTGTGTGAAG	r.736_741delinsugugugugaag, p.(Pro246Cysfs*3)	frameshift	5		3 (3)
c.794del	p.(Asn265Ilefs*42)	frameshift	[5]		1 (1)
c.904_911del	p.(Val302Thrfs*4)	frameshift	[5]		1 (1)
c.1020_1021del	p.(Arg341Alafs*23)	frameshift	[5]		1 (1)
c.1164del	p.(His388Glnfs*10)	frameshift	[5]	1 (0)	
c.1169_1170ins(20)	p.?	frameshift	5	1(1)	
c.1221del	p.(Thr408Leufs*40)	frameshift	5		1 (1)
c.1306dup	p.(Ser436Lysfs*22)	frameshift	5	1 (0)	
c.1486del	p.(His496Thrfs*99)	frameshift	[5]	1 (1)	
c.1500del	p.(Val501Trpfs*94)	frameshift	[5]	2 (1)	
c.1571dup	p.(Gly525Argfs*17)	frameshift	[5]	1 (1)	

PMS2 variant[a]	Change at RNA and/ or protein level[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
c.1579del	p.(Arg527Glyfs*68)	frameshift	[5]		1 (1)
c.1730dup	p.(Arg578Alafs*3)	frameshift	5		2 (2)
c.1768del	p.(Ile590Phefs*5)	frameshift	5	1 (1)	
c.1831dup	p.(Ile611Asnfs*2)	frameshift	5		2 (2)
c.2117del	r.2117del; p.(Lys706Serfs*19)	frameshift	[5]		1 (1)
c.2361_2364del	p.(Phe788Cysfs*2)	frameshift	5		1 (1)
c.137G>T	r.137g>u, p.(Ser46Ile)	missense	4	5 (5)	10 (10)
c.319C>T	r.[c>u, 301_353del, 251_353del], p.([Arg107Trp, ?, ?])	missense	[3]		1 (1)
c.505C>G	p.(Arg169Gly)	missense	[3]		1 (1)
c.614A>C	r.614a>c, p.(Gln205Pro)	missense	3		1 (1)
c.812G>T	p.(Gly271Val)	missense	[3]	1 (1)	
c.917T>A	p.(Val306Glu)	missense	[3]	1 (1)	
c.2113G>A	p.(Glu705Lys)	missense	3		1 (1)
c.2249G>A	p.(Gly750Asp)	missense	3		2 (2)
c.2444C>T	r.2444c>u, p.(Ser815Leu)	missense	3	1 (1)	
c.2531C>A	p.(Pro844His)	missense	[3]	1 (1)	
c.1A>G	p.?	variant in initiation codon	4		3 (3)
c.1A>T	p.?	variant in initiation codon	[4]		1 (1)
c.24-2A>G	p.?	canonical splice variant	[4]	1 (1)	
c.251-2A>C	p.?	canonical splice variant	[4]		1 (1)
c.803+2T>G	p.?	canonical splice variant	[4]		1 (1)

PMS2 variant[a]	Change at RNA and/ or protein level[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
c.804-2A>G	p.?	canonical splice variant	[4]		1 (1)
c.989-1G>T	p.?	canonical splice variant	5	1 (1)	
c.2007-2A>G	p.?	canonical splice variant	[4]	4 (1)	
c.2174+1G>A	p.?	canonical splice variant	5	1 (1)	1 (1)
c.2445+1G>T	r.2445_2446ins2445+1_2445+85, p.?	canonical splice variant	[4]	1 (1)	
c.825A>G	r.804_825del, p.(Ile269Alafs*31)	exonic splice variant	[3]		1 (1)
c.903G>T	r.804_903del; p.(Tyr268*)	exonic splice variant	4		1 (0)
c.24-12_107delinsAAAT	r.24_163del, p.(Ser8Argfs*5)	genomic deletion across canonical splice acceptor, resulting in skip of exon 2	5		2 (1)
genomic deletion including exon 1		large genomic deletion	5		1 (0)
genomic deletion including exon 7		large genomic deletion	5	3 (2)	
genomic deletion including exon 8		large genomic deletion	5		1 (1)
genomic deletion including exon 10		large genomic deletion	5		4(4)
genomic deletion whole gene (exons 1-15)		large genomic deletion	5		2 (2)
genomic deletion including exons 1-11		large genomic deletion	[5]		1 (1)

PMS2 variant[a]	Change at RNA and/ or protein level[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
genomic deletion including exons 5-15		large genomic deletion	5		1 (1)
genomic deletion including exons 5-7		large genomic deletion	5		1 (1)
genomic deletion including exons 6-15		large genomic deletion	[5]	1 (0)	
genomic deletion including exon 7-8		large genomic deletion (in frame)	[4]		2 (1)
genomic deletion including exon 8-9		large genomic deletion	[5]		1 (1)
genomic deletion including exon 9-15		large genomic deletion	5		1 (0)
genomic deletion including exons 12-14		large genomic deletion	[5]	1 (1)	
genomic deletion including exons 13-15		large genomic deletion	[5]		1 (1)
genomic deletion including exons 14-15		large genomic deletion	[5]	1 (1)	1 (1)
mutation(s) not identified				2 (0)	1 (0)

[a.] Variant nomenclature according to HGVS guidelines (http://varnomen.hgvs.org/) with reference to NM_000535.5 for *PMS2* except for the large deletions or duplications. Large deletions and duplications were in some cases detected with the older MLPA kit P008 (MRC Holland) that lacks reliable probes for *PMS2* exons 3, 4, 12-15. Therefore, the exact range of exon deletions was not always established. Although for some large deletions the breakpoints have been characterized, we did not include this information. [b.] As recommended by HGVS, protein changes are presented in parentheses (predicted consequences, i.e. without experimental evidence from protein sequence analysis); RNA changes are provided if experimental RNA analyses are performed (information on RNA analysis extracted from supplemental tables of Van der Klift *et al*. 2015 *Mol Genet Genomic Med 3(4):327-45,* and van der Klift *et al.* 2016 *Hum Mutat 37(11):1162-1179*). [c.] Clinical variant class as reported on https://insight-database.org/variants/PMS2, last accessed on July 14th, 2018; 5 = pathogenic, 4 = likely pathogenic, 3 = variant of uncertain significance. Variants not present or present but not yet classified in the InSiGHT database were classified by us using guidelines provided by https://www.insight-group.org/criteria/. Suggested classes are given in square brackets. Nonsense and frameshift mutations, including large genomic deletions, were classified as pathogenic (class 5). Variants in the initiation codon, canonical splice variants and large in-frame genomic deletions were classified as likely pathogenic (class 4). Information on the class 3 variants that could not be classified *a priori* as (likely) pathogenic (the missense variants and the exonic splice variant) is provided in supplemental table 3.

Supplemental table 2. *MSH6* variants

MSH6 variant[a]	(predicted) protein variant[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
c.642C>G	p.(Tyr214*)	nonsense	5		1 (1)
c.892C>T	p.(Arg298*)	nonsense	5	1 (1)	
c.2731C>T	p.(Arg911*)	nonsense	5	1 (1)	
c.2815C>T	p.(Gln939*)	nonsense	5		1 (1)
c.3020G>A	p.(Trp1007*)	nonsense	5		1 (1)
c.2653A>T	p.(Lys885*)	nonsense	[5]	1 (1)	
c.3450_3453del	p.(Ala1151*)	nonsense	[5]	1 (1)	
c.651dup	p.(Lys218*)	frameshift	5		1 (1)
c.1421_1422dup	p.(Gln475Cysfs*7)	frameshift	5		1 (1)
c.1596dup	p.(Glu533*)	frameshift	5		1 (1)
c.691del	p.(Val231Tyrfs*15)	frameshift	[5]	1 (1)	
c.1634_1635del	p.(Lys545Argfs*17)	frameshift	[5]		1 (1)
c.1800_1813dup	p.(Thr605Ilefs*10)	frameshift	[5]	1 (1)	
c.1998dup	p.(Asp667*)	frameshift	[5]		1 (1)
c.3261del	p.(Phe1088Serfs*2)	frameshift	5	1 (0)	2 (2)
c.3514dup	p.(Arg1172Lysfs*5)	frameshift	5		1 (1)
c.3482_3510del	p.(Pro1161Argfs*2)	frameshift	[5]	1 (1)	
c.3609_3612del	p.(His1203Glnfs*12)	frameshift	5		1 (1)
c.3635dup	p.(Asp1213Glyfs*2)	frameshift	5	1 (1)	
c.3939_3957dup	p.(Ala1320Serfs*5)	frameshift	5		1 (1)
c.3957dup	p.(Ala1320Serfs*5)	frameshift	5		1 (1)
c.3959_3962del	p.(Ala1320Glufs*6)	frameshift	5		1 (0)
c.3984_3987dup	p.(Leu1330Valfs*12)	frameshift	5		1 (0)
c.1763_1771dup	p.(His588_Pro590dup)	in-frame duplication	[3]	1 (1)	
c.2561_2563del	p.(Lys854del)	In-frame deletion	[3]		1 (1)

Supplemental table 2. *MSH6* variants

MSH6 variant[a]	(predicted) protein variant[b]	type of variant	classification[c]	Number of homozygous CMMRD index cases (families included in risk analysis)	Number of compound heterozygous CMMRD index cases (families included in risk analysis)
c.3386_3388del	p.(Cys1129_Val1130delinsLeu)	in-frame deletion	3	1 (1)	
c.2098C>T	p.(Leu700Phe)	missense	3		1 (1)
c.1196C>T	p.(Pro399Leu)	missense	[3]		1 (1)
c.2061T>G	p.(Cys687Trp)	missense	[3]		1 (1)
c.2087T>C / c.3163G>A (on one allele)	p.(Ile696Thr)/p.(Ala1055Thr)	missense/missense	both 3		1 (1)
c.2216C>T	p.(Thr739Ile)	missense	[3]	1 (1)	
c.3226C>T	r.3226c>u, p.(Arg1076Cys)	missense	4		2 (2)
c.3725G>A	p.(Arg1242His)	missense	[3]	1 (1)	
c.458-1G>A	p.?	canonical splice variant	4		1 (1)
c.3801+1_3801+5del	r.3647_3801del, p.(Arg1217Metfs*6)	canonical splice variant	[4]		1 (1)
c.3991C>T	r.[3991c>u, 3802_4001del], p.([Arg1331*, Ala1268Glyfs*6])	nonsense + exonic splice variant (partial skip exon 9)	5		1 (1)

[a.] Variant nomenclature according to HGVS guidelines (http://varnomen.hgvs.org/) with reference to NM_000179.2 for *MSH6*. [b.] As recommended by HGVS, protein changes are presented in parentheses (predicted consequences, i.e. without experimental evidence from protein sequence analysis); RNA changes are provided if experimental RNA analyses are performed. RNA analysis reports for NM_000179.2: c.3226C>T in Thompson *et al.* 2013 *Hum Mutat 34(1): 200–209*, for NM_000179.2: c.3991C>T in *Plaschke et al. 2006 Eur J Hum Genet 14(5):561-566.* For NM_000179.2: c.3801+1_3801+5del, unpublished RNA analysis data shows a skip of exon 8 and diminished expression of the mutant transcript through nonsense-mediated mRNA decay (NMD); absence of normal mRNA transcribed from mutated allele not tested therefore classified by us as a likely pathogenic variant class 4 (personal communication HM van der Klift). [c.] Clinical variant class as reported on https://insight-database.org/variants/ MSH6, last accessed on July 14th, 2018; 5 = pathogenic, 4 = likely pathogenic, 3 = variant of uncertain significance. Variants not present or present but not yet classified in the InSiGHT database were classified by us using guidelines provided by https://www.insight-group.org/criteria/. Suggested classes are given in square brackets. Nonsense and frameshift mutations were classified as pathogenic (class 5). Information on the class 3 variants that could not be classified *a priori* as (likely) pathogenic (the missense variants and the small in frame deletion or duplication), is provided in supplemental table 4.

Supplemental table 3. *PMS2* variants of uncertain significance

PMS2 variant[a]	Type of variant	homozygous/compound heterozygous (index included in this study)	CMMRD phenotype or reference describing CMMRD phenotype
c.319C>T r.[c>u, 301_353del, 251_353del] p.([Arg107Trp, ?]) (exon 4)	missense (+altered expression ratio of transcripts)	*in trans* with genomic deletion of exon 10	No information.
c.505C>G p.(Arg169Gly) (exon 5)	missense	*in trans* with c.1831dup p.(Ile611Asnfs*2)	Mork *et al*. Fam Cancer 2016:15(4):587-591
c.614A>C r.614a>c p.(Gln205Pro) (exon 6)	missense	*in trans* with c.1A>G	Senter *et al*. Gastroenterology 2008:135(2):419-28
c.812G>T p.(Gly271Val) (exon 8)	missense	homozygous	Kruger *et al*. Eur J of Hum Genet 2008:16: 62–72
c.825A>G r.804_825del p.(Ile269Alafs*31) (exon 8)	exonic splice variant	*in trans* with c.325dup	Johannesma *et al*. Clin Genet 2011:80: 243–255
c.917T>A p.(Val306Glu) (exon 9)	missense	homozygous	Two siblings with a CMMRD phenotype, details available upon request.
c.2113G>A p.(Glu705Lys) (exon 12)	missense	*in trans* with genomic deletion of exon 7-8	Lavoine *et al*. J Med Genet 2015: 52(11):770-8
c.2249G>A p.(Gly750Asp) (exon 13)	missense	two families: *in trans* with whole gene deletion (Senter 2008) *in trans* with genomic deletion of exon 10 (Lavoine 2015)	Senter *et al*. Gastroenterology 2008:135(2):419-28 Lavoine *et al*. J Med Genet 2015: 52(11):770-8
c.2444C>T r.2444c>u p.(Ser815Leu) (exon 14)	missense	homozygous	Suerink *et al*. Clin Genet 2018:93(1):134-137
c.2531C>A p.(Pro844His) (exon 15)	missense	homozygous	Lavoine *et al*. J Med Genet 2015:52(11):770-8

[a.] Variant nomenclature according to HGVS guidelines (http://varnomen.hgvs.org/) with reference to NM_000535.5 for *PMS2*.

Supplemental table 4. MSH6 variants of uncertain significance

MSH6 variant[a]	Type of variant	homozygous/compound heterozygous (index included in this study)	CMMRD phenotype or reference describing CMMRD phenotype
c.1196C>T p.(Pro399Leu) (exon 4)	missense	*in trans* with c.2061T>G p.(Cys687Trp)	Patient with a CMMRD phenotype and loss of MSH6 expression in cancer and normal tissue. Further information available upon request.
c.1763_1771dup p.(His588_Pro590dup) (exon 4)	in-frame duplication of 3 amino acids	homozygous	Lavoine *et al*. J Med Genet 2015:52(11):770-8
c.2061T>G p.(Cys687Trp) (exon 4)	missense	*in trans* with c.1196C>T p.(Pro399Leu)	see c.1196C>T
c.2087T>C/ c.3163G>A (on one allele) p.(Ile696Thr)/p. (Ala1055Thr)	missense/missense	*in trans* with c.2098C>T	Patient with a CMMRD phenotype at age 30, details available upon request.
c.2098C>T p.(Leu700Phe)	missense	*in trans* with c.2087T>C/ c.3163G>A (on one allele)	See c.2087T>C/ c.3163G>A
c.2216C>T p.(Thr739Ile) (exon 4)	missense	homozygous	Lavoine *et al*. J Med Genet 2015:52(11):770-8
c.2561_2563del p.(Lys854del) (exon 4)	in-frame deletion of 1 amino acid	*in trans* with c.3261dup p.(Phe1088Serfs*2)	Bougeard *et al*. Fam Cancer 2014:13(1):131-5 Lavoine *et al*. J Med Genet 2015:52(11):770-8
c.3386_3388del p.(Cys1129_ Val1130delinsLeu) (exon 5)	in-frame deletion, 2 amino acids replaced by another amino acid	homozygous	Menko *et al*. Fam Cancer 2004:3(2):123-7
c.3725G>A p.(Arg1242His) (exon 8)	missense	homozygous	Two siblings with a CMMRD phenotype and loss of MSH6 expression in the tumor and normal tissue. Functional testing as described by Bodo *et al*.[b] showed *ex vivo* microsatellite instability and tolerance to methylation. Further details available upon request.

[a.] Variant nomenclature according to HGVS guidelines (http://varnomen.hgvs.org/) with reference to NM_000179.2 for *MSH6*
[b.] Bodo *et al. 2015 Gastroenterology 2015 149(4):1017–1029*

Supplemental table 5. Frequency of other cancers

Cancer type	PMS2 (n=93)	MSH6 (n=34)
Leukaemia	**10**	**3**
Acute	2	2
Chronic	1	1
Not specified	7	0
Lymphoma	**0**	**2**
Gynaecological	**3**	**2**
Ovarian	1	0
Cervical	2	2
Prostate	**9**	**2**
Testicular	**0**	**1**
Respiratory tract	**17**	4
Lung	13	3
Upper airway/throat	4	1
Gastrointestinal tract	**16**	**7**
Biliary tract	3	0
Hepatic	1	1
Pancreatic	1	0
Duodenal	2	2
Stomach	6	2
Oesophageal	3	1
Not further specified	0	1
Urinary tract	**2**	**4**
Kidney	1	0
Bladder/ureters	1	4
Breast	**14**	**2**
Eye	**0**	**1**
Melanoma	**4**	**1**
Mesothelioma	**0**	**1**
Brain	**4**	**1**
Thyroid	**1**	**1**
Bone	**1**	**1**
Rhabdomyosarcoma	**1**	**0**
Teratoma	**1**	**0**
Mullerian tumor	**1**	**0**
Tumor of unspecified site	**9**	**1**

SUPPLEMENTALS STATISTICAL METHODS

Colorectal cancer (CRC) risk estimates are corrected by the presence of competing risks given by death and other cancer diagnoses, to account for the realistic possibility of the studied mutation affecting other cancer incidences and death. In general, the observed data in a competing risk setting is given by the failure time *T*, and the cause of failure *D (D=1,...k).* In our case, we denote by *k* the cause of interest, CRC, and the CRC risk at age *t* is estimated by the cumulative incidence:

$$I_k(t|x) = \int_0^t h_k(s|x_i)S(s|x_i)ds \quad (1)$$

In this expression $h_k(t|x) = lim_{\Delta t \to 0} \frac{P(t \leq T < t+\Delta t, D=k|T \geq}{\Delta t}$ is the cause-specific hazard function, the hazard of failing from a given cause (CRC in our case) in the presence of the competing events (death and other cancer diagnosis) and *x* is the covariate sex. h_k is estimated using proportional hazard regression:

$$h_k(t|x) = h_{k,0}(t)\exp(\beta x) \quad (2)$$

In this equation is the baseline cause-specific hazard of cause *k* (CRC) and β is the effect of sex on cause *k*. To deal with the missing carrier status of some of the included individuals, we perform weighted regression, by including mutation probabilities as weights in the score function:

$$U_w(\beta) = \sum_{i=1}^{n} w_i \left[x_i - \frac{\sum_{j \epsilon R_i} w_j x_j \exp(x_j \beta)}{\sum_{j \epsilon R_i} w_j \exp(x_j \beta)} \right] \quad (3)$$

Analytical weight for individual *j, wj =* is given by the kinship coefficient between individual *j* and the closest family member with observed mutation. This probability is always positive for all the individuals in the studied cohort given the design based on the identification of at least one member carrying a biallelic mutation in each included family.

Once the cause-specific hazard is estimated using expressions (2) and (3), the cumulative cause-specific hazard can be calculated as $\Lambda_k(t|x) = \int_0^t h_k(s|x_i)ds$ and the marginal survival function, $S(t|x) = exp\left(-\sum_{k=1}^{K} \Lambda_k(t)\right)$ which is plugged in expression (1) to obtain the cumulative incidence of interest.

Confidence intervals (CI) were obtained by bootstrapping at family level (1,000 repetitions) to account for possible dependencies between family members.

REFERENCES

1. Putter H, Fiocco M and Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. Statistics in Medicine. 2007; 26: 2389-2430.