

# The Swedish Family-Cancer Database: Update, Application to Colorectal Cancer and Clinical Relevance

Kari Hemminki<sup>1,2</sup>, Charlotta Granström<sup>2</sup>, Bowang Chen<sup>1</sup>

<sup>1</sup>Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany;

<sup>2</sup>Department of Biosciences at Novum, Karolinska Institute, Huddinge, Sweden

**Key words:** familial cancers, offspring, familial risk, colorectal cancer, recessive effects, clinical counselling

**Corresponding author:** Kari Hemminki, Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany, e-mail: k.hemminki@dkfz.de

Submitted: 21 December 2004

Accepted: 10 January 2005

## Abstract

The Swedish Family-Cancer Database has been used for almost 10 years in the study of familial risks at all common sites. In the present paper we describe some main features of version VI of this Database, assembled in 2004. This update included all Swedes born in 1932 and later (offspring) with their biological parents, a total of 10.5 million individuals. Cancer cases were retrieved from the Swedish Cancer Registry from 1958-2002, including over 1.2 million first and multiple primary cancers and *in situ* tumours. Compared to previous versions, only 6.0% of deceased offspring with a cancer diagnosis lack any parental information. We show one application of the Database in the study of familial risks in colorectal adenocarcinoma, with defined age-group and anatomic site specific analyses. Familial standardized incidence ratios (SIRs) were determined for offspring when parents or sibling were diagnosed with colon or rectal cancer. As a novel finding it was shown that risks for siblings were higher than those for offspring of affected parents. The excess risk was limited to colon cancer and particularly to right-sided colon cancer. The SIRs for colon cancer in age matched populations were 2.58 when parents were probands and 3.81 when siblings were probands; for right-sided colon cancer the SIRs were 3.66 and 7.53, respectively. Thus the familial excess (SIR-1.00) was more than two fold higher for right-sided colon cancer. Colon and rectal cancers appeared to be distinguished between high-penetrant and recessive conditions that only affect the colon, whereas low-penetrant familial effects are shared by the two sites. Epidemiological studies can be used to generate clinical estimates for familial risk, conditioned on numbers of affected family members and their ages of onset. Useful risk estimates have been developed for familial breast and prostate cancers. Reliable risk estimates for other cancers should also be seriously considered for routine clinical recommendations, because practically all cancers show a familial effect and the risks are high for some of the rare neoplasms. The implementation of a unified management plan for familial cancers at large will be a major challenge to the clinical genetic counselling community.

## Introduction

The majority of cancers are sporadic and some 1-5% are due to single-gene, dominant traits [1]. There are no estimates on the contribution of polygenic or

recessive conditions on cancer, but twin studies suggest that such effects are important [2]. Familial aggregation of cancer may be due to environmental factors shared by family members or due to shared genes.

**Table 1. Number of cancer notifications for first and multiple primary invasive and in situ cancers in the Family-Cancer Database, 1958-2002**

	First primary	Multiple primary	<i>In situ</i>	All
Father	315,921	37,379	30,057	383,357
Mother	270,039	33,676	64,203	367,918
Father/son	65,269	3,832	6,982	76,083
Mother/daughter	97,442	7,658	114,976	220,076
Son	24,230	1,105	1,752	27,087
Daughter	23,075	1,614	18,671	43,360
All	795,976	85,264	236,641	1,117,881

Observation of familial aggregation has been important for the understanding of cancer aetiology, for clinical decisions and counselling and for identification of cancer-related genes. Our comparisons of familial cancer risks between spouses, and between parents and offspring suggest that the known environmental factors explain familial aggregation for lung and gastric cancer and for melanoma, but only to a limited degree. At the remaining sites inherited factors appear to be the main cause [3].

Familial clustering of cancer has been studied through clinical identification of probands and multiple affected family members [1, 4]. This approach has been very productive also in terms of understanding cancer genetics. Many forms of cancer in which a single gene poses a high risk have been identified. The disadvantages include difficulties in obtaining large numbers of cases and in securing unbiased risk estimates. Also estimation of risks at sites other than the index site has been cumbersome. Clinical observation probably works for dominant diseases where risks are between 10 and 100, or more. For recessive conditions, it is less sensitive, and most results on recessive conditions have come from an isolated population with high rates of consanguineous marriage. Population geneticists have raised questions about a relatively small number of known human recessive syndromes. In species of experimental animals recessive traits predominate as opposed to humans where dominant traits are more common [5]. It is not excluded that this is an observation bias because of difficulties in identifying a recessive pattern.

A major problem in the global literature on familial cancer, which overwhelmingly consists of case-control studies, is the possible inaccuracy of data on cancer in family members who had died a long time before

the study. In hardly any study, except those based on registers, are cancer diagnoses confirmed both for the cases and the probands. The false reporting for internal cancers may be as high as 50%. This level of inaccuracy may cause a severe bias in the derived risk estimates, with a tendency to report exaggerated risks, in the most commonly used case-control studies [6]. A second problem is that all studies have been small and it has been difficult to obtain reliable information on familial aggregation of rare cancers. Overall, formal epidemiological studies have had little impact on defining new familial traits. They have been hypothesis either generating or quantifying the known risks. The Utah Population Database has been used in quantifying known risks [7] or in gene identification. The largest dataset on familial cancer is the Swedish Family-Cancer Database. In the present article we report on 2004 update of this resource, and show an application of the Database on colorectal cancer. Additionally we discuss the application of epidemiologically derived risk estimates in clinical genetic counselling.

## The Family-Cancer Database as of 2004

Statistics Sweden created a family database, "Second Generation Register" in 1995 [8]. After a few expansions, it covered offspring born after 1931 with their parents, and it was renamed to "Multigeneration Register". We have linked the Second Generation Register to the Swedish Cancer Registry (started in 1958) to make the Family-Cancer Database in five expanded versions in 1996, 1997, 1999, 2000 and 2002. The number of cancers in the second generation increased from 20,000 in 1996 to 158,000 in 2000; in the parental generation the increase was from 500,000 to 602,663 invasive cancers. In the most recent update, version VI, completed in 2004, the number of invasive cancers in the 0- to 70-year-old

Table 2. Number of offspring in the Family-Cancer Database in 1958-2002

	Total no. of offspring	Offspring with cancer		Offspring linked to parent		Offspring with cancer not linked to parent	
	n	n	%	n	%	n	%
All	7,400,436	224,225	3.0	7,240,222	97.8	26,254	0.4
Living offspring	7,100,978	144,602	2.0	7,010,092	98.6	8,936	0.1
Deceased offspring	289,458	79,623	27.5	230,130	79.5	17,318	6.0

offspring generation has increased to 224,000. In the Family-Cancer Database all data are organized in child-mother-father triplets; the parents have been registered at the time of birth of the child, allowing tracking of "biological" parents in spite of divorce and remarriage. The national personal identification code (personnummer) has been deleted from the Database.

Version VI of the Database identified a total of 10.5 million persons, of whom 1,075,597 patients were diagnosed with any tumour (Table 1). These included 795,976 first primary cancers, 85,264 multiple primaries and 236,641 *in situ* cancers, distributed between the generations. The annual accumulation of cancer cases to the parental and offspring generation in the Family-Cancer Database is shown in Fig. 1. The offspring

generation has far fewer cases than the parental one but the increase among the offspring generation is steep due to the advancing age. The 'All' curve is the sum of the three other mutually exclusive curves.

The linkage of offspring to their parents was partially incomplete among the deceased individuals, particularly among those who were born in the 1930s and who died before the 1990s. According to Table 2, 3.0% of the 7.4 million offspring had cancer, and 97.8% had a link to at least one parent. The linkage was 98.6% among those alive as of the end of 2002. Among the deceased 289,458 individuals, the linkage existed among 79.5%. Parental information was lacking on 17,318 deceased offspring, who had been diagnosed with cancer; this was 6.0% of the deceased offspring and implied that

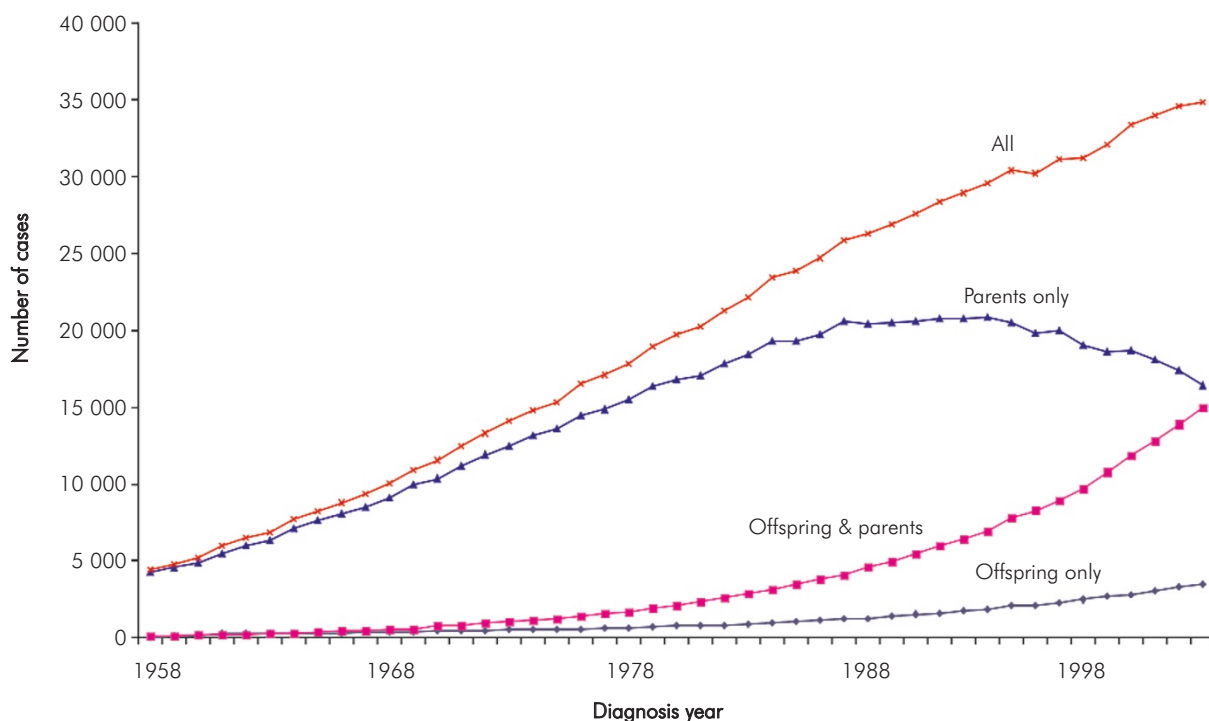


Fig. 1. Annual accumulation of cancer cases to the parental and offspring generation of the Family-Cancer Database 2002

Table 3. SIR for offspring CRC depending on the CRC in probands

Familial cancer site	Parent only				Sibling only				Parent and Sibling			
	O <sup>1)</sup>	E	SIR	95%CI	O <sup>1)</sup>	E	SIR	95%CI	O	E	SIR	95%CI
Colon	448	239.40	<b>1.87</b>	1.70 2.05	68	21.70	<b>3.13</b>	2.43 3.97	13	1.50	<b>8.73</b>	4.63 14.97
Rectum	227	141.30	<b>1.61</b>	1.40 1.83	30	14.10	<b>2.13</b>	1.44 3.04	1	0.40	<b>2.24</b>	0.00 12.81
Colorectum	664	376.30	<b>1.76</b>	1.63 1.90	93	34.30	<b>2.71</b>	2.19 3.32	17	3.30	<b>5.08</b>	2.95 8.16

Bold type, 95% CI does not include 1.00.

<sup>1)</sup>Some observed numbers do not add up because in addition to the right- and left-sided colon cancer, multiple and unspecified colon tumours were included.

0.4% of all offspring with cancer had no links to parents. Offspring who died before 1960 are completely missing from the Database. The records have been practically complete since 1990 when the overwhelming majority of familial cancers were recorded [9].

A relevant question is the generalisability of the Swedish results to the mankind at large. We believe that the results are generalisable to the world population which has approximately the same level of cancer, i.e. developed countries. However, considering that for some familial risks "gene-environment" interactions may play a role, some caution needs to be exercised when making extrapolations to populations whose cancer incidence is appreciably different from that in the Nordic countries. Consistently, the estimates of familial risk for breast and colorectal cancer between various populations have not shown evidence of population-dependent variation [10, 11]. Moreover, in a series of studies on cancer risks among the Swedish immigrant populations, no differences in cancer risk have been found between the second generation immigrants and Swedes [12-14]. According to a recent review, racial differences are small in complex diseases, including cancer [15].

### Application of the Database: colorectal cancer

A family history of colorectal cancer (CRC) is a risk factor for CRC, the contribution of which to all CRC ranges, depending on the definition, from 7% for co-aggregation in nuclear families to 35% for clustering among twins [2, 16, 17]. The most common Mendelian condition, hereditary nonpolyposis colorectal cancer (HNPCC) features a relatively early onset (average 45 years) and preferential involvement of the right-sided colon [17]. It affects also other sites, such as the endometrium, pancreas, ovary, stomach and upper urothelial tract. It has been estimated to

account for 3% of all CRC in Finland [18, 19] and around 1% in Sweden and the USA [20-22], the variation depending on many parameters, including the age structure of the studied population and the existence of founder mutations. Familial adenomatous polyposis (FAP), Peutz-Jeghers syndrome, familial juvenile polyposis and Cowden's disease are other syndromes in which CRC may be a manifestation [17, 23, 24]. All these Mendelian conditions follow a dominant mode of inheritance. Recent data indicate that recessive germline mutations in the *MYH* gene may also increase the risk of CRC [25-28]. The suggested role for the *BML* in CRC has not been confirmed [29, 30]. Even some other low-penetrant genes may also contribute to familial risk [31-34]. The pending questions regarding familial CRC are characteristics of the conditions which cannot be accounted for by the known syndromes, including their mode of inheritance. Furthermore, any unique familial risks at anatomic subsites of CRC are of interest. For the interpretation of familial risks, it is important to estimate the contribution of environmental sharing to familial clustering. We have recently concluded, by comparing risks for CRC between various family members, that the familial clustering of CRC is mainly due to heritable causes [35].

We use the Family-Cancer Database, version V, to examine some outstanding questions about familial CRC. We analyse the risk of colorectal adenocarcinoma in offspring whose parents or siblings present with these cancers, by adhering specifically to adenocarcinoma and considering colonic subsites. Standardized incidence ratios (SIRs) were calculated for offspring CRC as the ratio of observed (O) to expected (E) number of cases. The expected numbers were calculated from 5-year-age-, sex-, period- (10-year bands), area- (county), socio-economic status standardized rates [36]. Confidence intervals (95% CI) were calculated assuming a Poisson distribution [36]. In the calculation of the CIs for sibling risks, the

Table 4. SIR for colorectal adenocarcinoma subsites in offspring by proband colorectal adenocarcinoma

Offspring	Proband	Proband subsite											
		Right-sided colon			Left-side colon			All colon			Rectum		
		O <sup>1)</sup>	SIR	95%CI	O <sup>1)</sup>	SIR	95%CI	O	SIR	95%CI	O	SIR	95%CI
Right-sided colon	Parent	70	<b>2.45</b>	1.91 3.10	44	<b>1.90</b>	1.38 2.55	128	<b>2.11</b>	1.76 2.51	42	<b>1.39</b>	1.00 1.88
	Sibling	16	<b>6.89</b>	3.93 11.21	12	<b>5.16</b>	2.65 9.03	29	<b>5.29</b>	3.54 7.60	6	1.68	0.60 3.68
	Parent and sibling	4	<b>34.62</b>	9.00 89.52	0			6	<b>16.04</b>	5.77 35.14	1	10.66	0.00 61.11
Left-side colon	Parent	51	<b>1.85</b>	1.38 2.44	48	<b>2.15</b>	1.58 2.85	113	<b>1.94</b>	1.60 2.34	52	<b>1.79</b>	1.33 2.34
	Sibling	10	<b>4.56</b>	2.17 8.41	4	1.81	0.47 4.68	15	<b>2.89</b>	1.61 4.77	9	<b>2.66</b>	1.20 5.06
	Parent and sibling	2	<b>18.54</b>	1.75 68.18	0			5	<b>14.00</b>	4.42 32.94	0		
All colon	Parent	147	<b>2.16</b>	1.82 2.53	109	<b>1.97</b>	1.62 2.38	292	<b>2.02</b>	1.80 2.27	122	<b>1.43</b>	1.19 1.71
	Sibling	29	<b>5.23</b>	3.50 7.52	16	<b>2.89</b>	1.65 4.70	50	<b>3.81</b>	2.83 5.03	19	<b>2.24</b>	1.34 3.50
	Parent and sibling	6	<b>21.94</b>	7.90 48.06				12	<b>13.39</b>	6.88 23.46	1	3.65	0.00 20.94
Rectum	Parent	69	<b>1.54</b>	1.20 1.95	58	<b>1.59</b>	1.21 2.05	156	<b>1.64</b>	1.39 1.92	84	<b>1.77</b>	1.41 2.19
	Sibling	5	1.38	0.43 3.24	9	<b>2.43</b>	1.10 4.64	18	<b>2.09</b>	1.23 3.31	11	1.96	0.97 3.51
	Parent and sibling	1	5.93	0.00 34.02	0			1	1.69	0.00 9.66	0		

Bold type, 95% CI does not include 1.00.

<sup>1)</sup>Some observed numbers do not add up because in addition to the right- and left-sided colon cancer, multiple and unspecified colon tumours were included.

dependence between the affected pairs was taken into consideration [37]. The SIRs for offspring and siblings when parent-only, sibling-only and parent and sibling together being probands of different cancers were calculated respectively.

### Familial risk of CRC in offspring

The Family-Cancer Database (version V) covered years 1958 to 2000 from the Swedish Cancer Registry and included 11,921 offspring and 80,196 parents with CRCs of adenocarcinoma histology. Because we only covered adenocarcinoma histology, the term CRC later in this study refers to this histology only. Among affected offspring (age 0-68 years), 664 had a parent, 93 a sibling and 17 both a parent and sibling diagnosed with CRC (Table 3). The familial risk was somewhat higher for colon than for rectal cancer but the difference was higher (SIR 3.13 for colon and 2.13 for rectal cancer) with a fraternal (sibling) proband compared to a parental proband (SIR 1.87 for colon and 1.61 for rectal cancer). When both a parent and a sibling were affected, only colon cancer was increased (SIR 8.73).

Age-specific SIRs of CRC in offspring by proband CRC are shown in Table 4. The SIR for colon cancer was 2.02 when parents were diagnosed with colon cancer. SIR for colon cancer in siblings was 3.81 and it was 13.39 when additionally a parent was affected. The SIR for concordant rectal cancer in offspring and parents was 1.77 and that among siblings was 1.96, of borderline significance. The highest SIRs were found for concordant right-sided colon cancer. The SIR for offspring of affected parents was 2.45, for siblings it was 6.89 and for families of affected parents and offspring it was 34.62. The SIRs for siblings were also vastly increased when the pairs presented with right- and left-sided colon cancer (5.16 and 4.56); however, no significant increase was found for siblings presenting with right-sided colon and rectal cancers.

The direct comparison of risks between offspring of affected parents and siblings requires that ages of the populations are similar, which is not the case in Table 3 and 4 because parents are of any age and all siblings are below the age of 69 years. When the parental age was limited to 68 years and younger, the SIR for colon

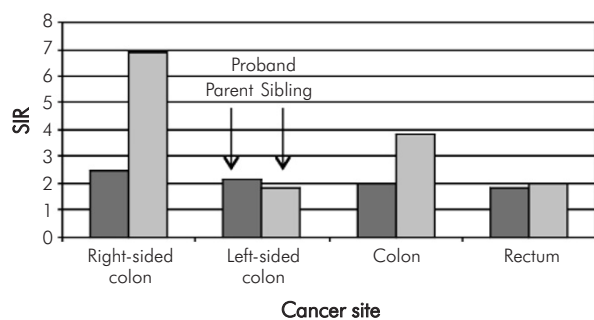


Fig. 2. Comparison of SIRs of parent and sibling proband in CRC subsites

cancer in offspring of affected parents increased to 2.58 (N=138, 95% CIs 2.17-3.05) compared to the unchanged sibling risk of 3.81. For right-sided colon cancer the risk in offspring of affected parents was 3.66 (35, 2.55-5.09); the sibling risk was somewhat increased from Table 4 because of another definition of the population to 7.53 (18, 4.46-11.93). For the whole colon cancer the familial excess risk (SIR minus 1.00) was 1.8 times higher for siblings (3.81-1.00=2.81) than for parents of affected parents (2.58-1.00=1.58).

The median diagnostic age showed small differences between the anatomic sites when parents were probands (data not shown). However, when parents were not affected, the median diagnostic age for sibling with concordant colon and rectal cancer was 54.3 and 55.9 years, respectively. For the 16 siblings with concordant right-sided colon cancer, the median age was 53.3 years, compared to 56.8 years in those 4 siblings with concordant left-sided colon cancer.

The notable difference in sibling risks for right-sided colon cancer compared to other concordant sites is illustrated in Fig. 2. By contrast, the SIRs for offspring of affected parents show only a minor difference. In order to search evidence for HNPCC in families of affected siblings, at least one of whom was diagnosed with right-sided colon cancer, we listed all first and second cancer in all their family members. In these 19 families a total of 53 offspring were found, in addition to those 38 with colon cancer. In one family of two siblings with right-sided colon cancer a third sibling was diagnosed with rectal cancer and the father was diagnosed with pancreatic cancer. This was the only family with a likely HNPCC diagnosis. However, the presentation of HNPCC-related tumours was much

more frequent in the 9 families where a parent and at least two siblings were diagnosed with CRC.

Effects of diagnostic age for offspring and parents are shown as contour plots in Fig. 3 for colon cancer and in Fig. 4 for rectal cancer. The SIR curves have a linear diagonal direction for colon cancer, indicating that the diagnostic age of both the offspring and the parents are inversely related to the risk. SIRs are >10 when both were diagnosed before the age of 50 years. For rectal cancer, the contour plots are nonlinear; the highest risk is among young offspring whose parents were diagnosed at ages around 50 years. Colon was further divided to right- and left-sided segments (data not shown). The plot for the right-sided colon resembled that for the whole colon but with steeper effect of age at ages below 50 years. The contour for the left-sided colon cancer resembled that for rectal cancer but the parental age maximum was 5 years lower.

### Interpretation of results for colon and rectal cancer

There has been a wealth of previous literature on familial risks in CRC. The review of Johns and Houlston covered literature until 1999 and it included 27 studies [11]. The relevant pooled estimates, given as relative risks of CRC, were 2.42 (95% CIs 2.20-2.65) for colon cancer in a first degree relative and 1.89 (1.62-2.21) for rectal cancer in a relative; the CRC risks were 2.26 (1.87-2.72) and 2.57 (2.19-3.02) for a parental and sibling proband with CRC, respectively. The authors discussed many aspects of validity of the referred data, including the need to verify cancer diagnoses of both the cases and their relatives, which has been achieved in a limited number of studies. The published results from the Swedish Family-Cancer Database, with fully medically verified diagnostic data, have not been essentially different [38]. However, when colon and rectal cancers were analysed separately, the SIRs for colon and rectal cancer by parental proband did not differ from each other and they were somewhat lower than the cited figures; on the other hand, the sibling risks were higher, particularly when colon cancer was involved [39]. The special features of the present study, in addition to the size, are focused on adenocarcinoma only, data on age of onset, separation of colonic subsites and distinction of probands. The latter condition allows modes of inheritance to be estimated.

The known cancer syndromes manifesting CRC are dominant Mendelian, except for the recently identified *MYH* genes exerting recessive effects on CRC [25]. Biallelic mutations in the *MYH* gene have

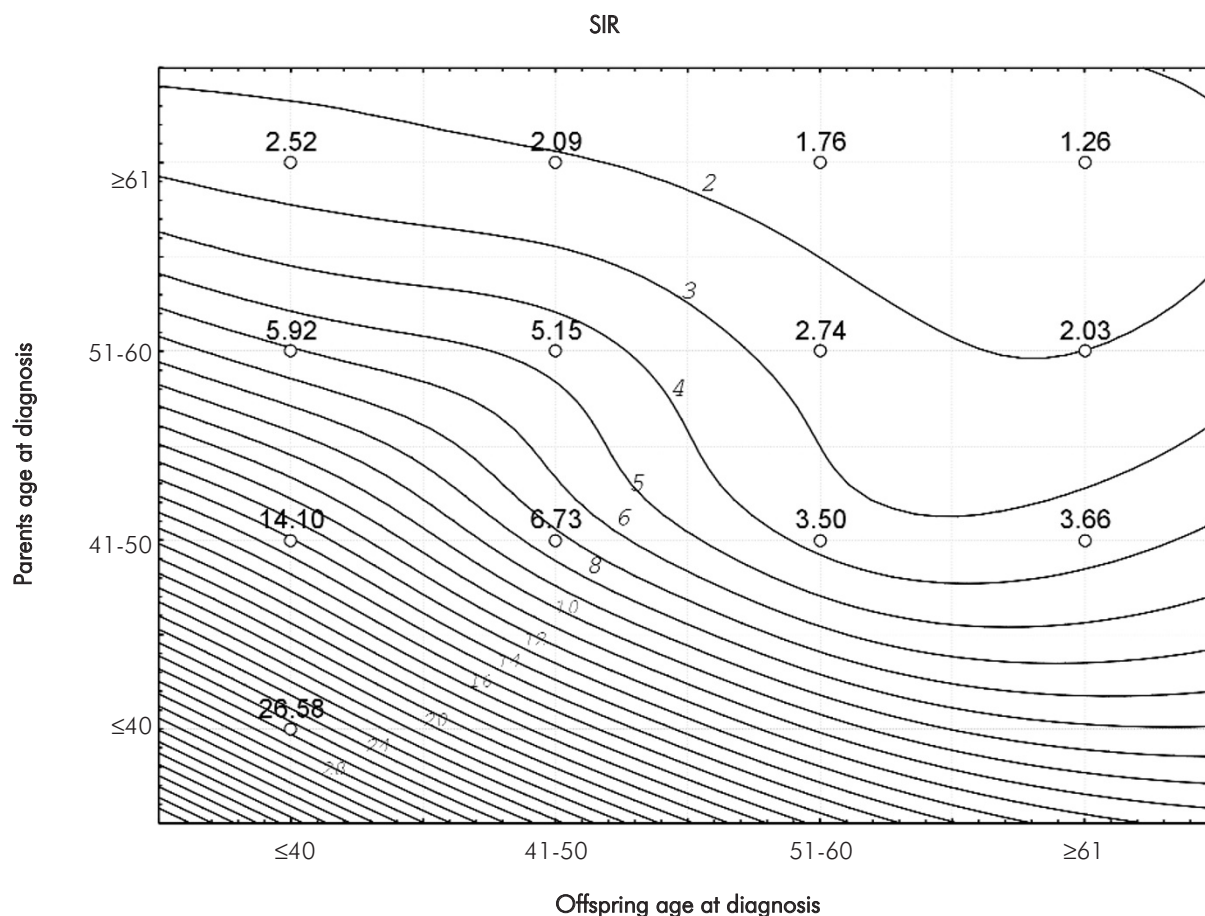


Fig. 3. SIR trend of colon cancer for offspring by the ages at diagnosis of parents and offspring

been found in about 0.5% of CRC patients but only 6 patients have been described in the literature so far [26, 28]. All except one of these patients presented with left-sided colon and rectal tumours. The important new finding is that risks for siblings were higher than those for offspring of affected parents but the excess risk was limited to colon cancer and particularly to right-sided colon cancer. The SIRs for colon cancer in age matched populations were 2.58 when parents were probands and 3.81 when siblings were probands; for right-sided colon cancer the SIRs were 3.66 and 7.53, respectively. Thus the familial excess (SIR-1.00) was more than two fold higher for right-sided colon cancer from sibling than parental probands (6.53 vs. 2.66). As our recent data find no evidence for an environmental contribution to sibling risk even for right-sided colon cancer [35], the results show the existence of a recessive inheritance at this colonic subsite. In order to exclude known causes of CRC, we examined the available data on first and second primary cancer among the affected siblings

and their family members. HNPCC was likely only in one family, in which 2 siblings were diagnosed with right-sided colon cancer, because another sibling was diagnosed with rectal cancer and the father with pancreatic cancer. If indeed a recessive inheritance has been found, it does not appear to cause a high risk of other tumours. Among the 8 sibling pairs with right-sided colon cancer, none of their other 20 siblings had cancer; only one affected individual was diagnosed with second cancer, which was in the cervix. Similarly, among 11 sibling pairs with right-sided and left-sided colon cancer, only one of 32 siblings was diagnosed with cancer, which was in the lung; one of the affected siblings was diagnosed with a second prostate cancer. Even considering the diagnosed tumours in parents of these 19 families of the affected siblings (8 with two right-sided colon cancer and 11 with pairs of right- and left-sided colon cancers), no evidence was found for a recognizable cancer syndrome, except the above family with a likely HNPCC. It should be pointed out that although previous studies,

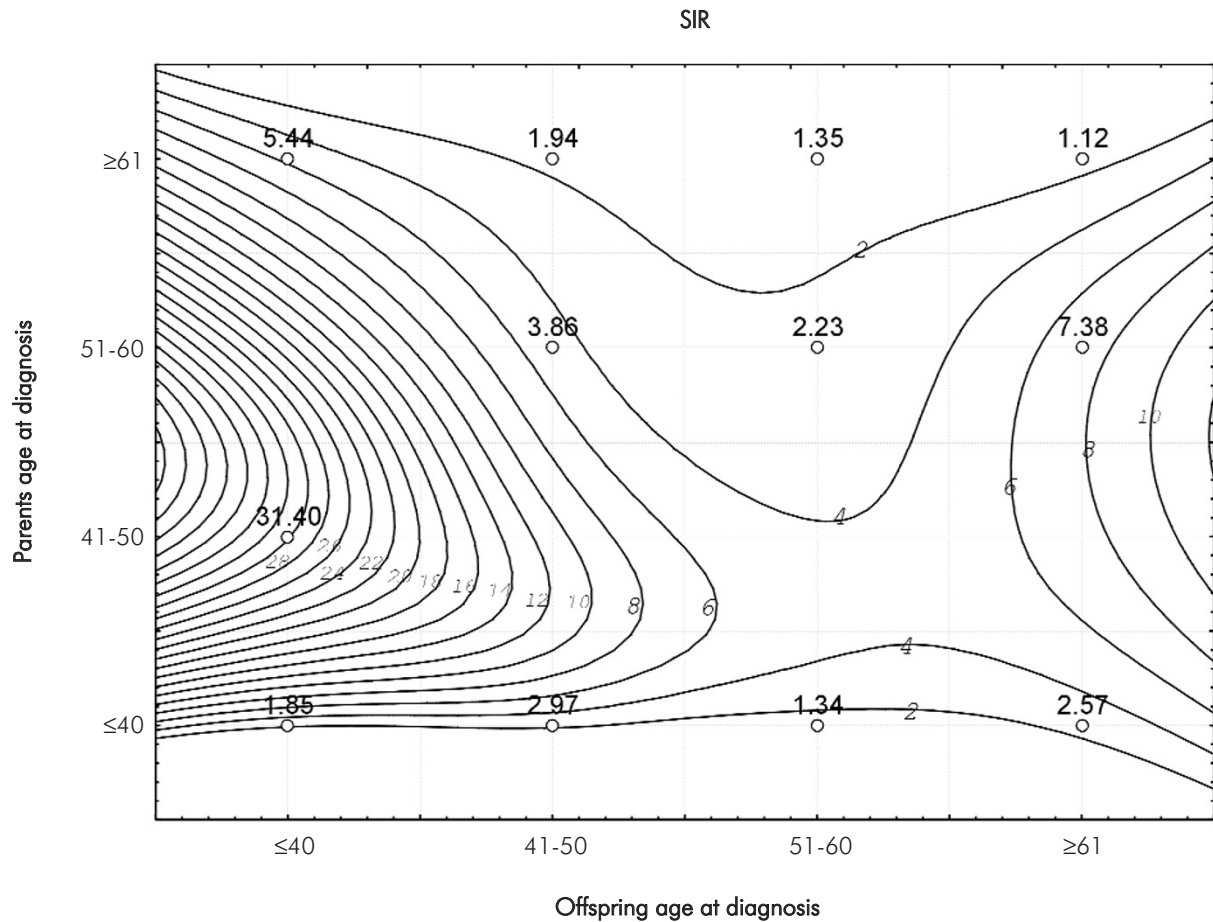


Fig. 4. SIR trend of rectal cancer for offspring by the ages at diagnosis of parents and offspring

including the meta-analysis by Johns and Houlston [11], have not found a difference in the sibling risk for CRC, the largest family study published outside Sweden did find a difference [40], and this Danish study was based on fully medically verified cases, like our present study, adding credibility to the finding.

We have reasons to believe that the detected evidence on a recessive inheritable cannot be explained by *MYH* mutations. Firstly, in the published studies only one of 6 patients was diagnosed with a right-sided colon tumour [26, 28]. Secondly, the reported frequency, of about 0.5%, appears to be less than the one noted in the present study. For the whole colon cancer the familial excess risk was 1.8 times higher for siblings than for parents of affected parents, suggesting that in this population of 0- to 68-year-old individuals close to half of the familial risk for colon cancer can be explained by recessive inheritance.

Cancers of the colon and rectum are often considered together as CRC because of biological and functional similarity [41, 42]. However, there is evidence that the colon and the rectum respond differently to the main risk factors, such as obesity and physical activity [43]. The previous data on the possible role of the family history have been inconclusive because of small numbers but, for example, HNPCC is known to have a predilection for proximal, right-sided colon cancer [17, 39, 41]. The present data show that the familial SIR for offspring of affected parents is not different between colon (SIR 2.02) and rectal (1.77) cancer. On the other hand, sibling risk is appreciably higher for the colon (3.81) than for the rectum (1.96, borderline significance). Moreover, familial triplets of two or more affected siblings in families of affected parents were only found for colon cancer (SIR 13.39). Thus colon and rectum appear to be distinguished between high-penetrant and recessive conditions that only affect the colon, whereas low-penetrant familial effects are shared by the two sites.



Effects of diagnostic age for offspring and parents showed different types of contour plots for colon and rectal cancer. The SIR curves followed a linear diagonal pattern for colon cancer, indicating that the diagnostic age of both the offspring and the parents were inversely related to the risk. SIRs are  $>10$  when both were diagnosed before the age of 50 years. For rectal cancer, the contour plots were nonlinear, the highest risk being found for young offspring whose parents were diagnosed at ages around 50 years. The plot for the right-sided colon resembled that for the whole colon but with steeper effect of age at ages below 50 years. The contour for the left-sided colon cancer resembled that for rectal cancer but the parental age maximum was 5 years lower. The other relevant new data relate to sibling risks. Particularly in colon cancer, the majority of the affected siblings presented with a relatively late onset disease and without an affected parent, thus suggesting involvement of low-penetrance or recessive genes unrelated to HNPCC. These non-HNPCC familial CRCs will probably be difficult to characterize genetically, because they may be heterogeneous with low penetrance, recessive mode and no strong association to other types of cancers.

### Clinical use of data on familial risks

The algorithms for clinical genetic counselling need to be based on reliable data on familial risks with consideration of the genetic data on possible underlying genes [44]. For cancer syndromes with identified genes, mutation testing may be offered. However, the known syndromes cover a small proportion of familial aggregation of cancer and counselling has to rely on epidemiologically derived risk estimates. Indeed, many risk estimation models were developed before any genes were identified, and for prostate cancer no established genes can yet be offered for testing. The two clinically-utilised models for breast cancer risk assessment were devised before the identification of the highly penetrant breast cancer susceptibility genes. The Gail model predicts a woman's risk for breast cancer, based on her individual risk factors, including a family history [45]. The model was developed using data from a large follow-up study. The Gail model has been evaluated in several settings and it has been found to give reasonably accurate predictions of risk [46]. The Claus model has been developed based on a case-control study and it has been useful for the estimation of heritable risks of breast cancer [47]. However, because neither of these models takes into account family history, hormonal factors and

benign breast disease comprehensively, both models have been found to systematically underestimate risk of developing cancer in women who attend a cancer genetics clinic [48]. A recent model which does take all these factors into account, the Tyrer-Cuzick model, appears to be more consistently accurate [49]. These are examples of how detailed epidemiological data provide relatively valid risk predictions in the absence of mutation analysis. Once mutational data become known, a further refinement of the risk assessment models must necessarily be done.

In low-penetrant cancers, familial aggregation is present, but Mendelian patterns cannot be ascertained, genetic mechanisms have not been worked out and the only types of available risk estimates are derived from epidemiological studies. In many cancers, different levels of familial risk can be discerned without aetiological explanation. For example, the Groningen Database on familial cancer, the Familial Cancer Database (<http://www.facd.info>) contains a large number of data on familial cancer clusters in which no specific genes have been identified, yet with a great importance to clinicians and other specialists who have to advise patients and their families [50]. Typically, prostate cancer displays high-risk and lower-risk familial clusterings, depending on the age of onset and family relationships. Clinical genetic counselling is currently based on these parameters. For some cancers, screening tests are available and they may be recommendable in familial cases, irrespective of whether the genetic background of the disease is known. However, for most familial cancers risk estimates are only derived from epidemiological studies of the kind proposed in the present application. Familial cancer clustering, without obvious heritability, poses a major challenge to current cancer risk assessment and management. Reliable determination of familial risks for cancer is important for clinical genetic counselling, but medically verified data on familial risks for many malignancies have been limited. The data on familial risk generated in population-based studies are helpful in implementing evidence-based guidelines for helping the general medical system to ascertain and refer even familial cancer clusters to cancer genetic professionals. Familial SIR and relative risks need to be translated to the individual-based risk estimates useful in cancer genetics counselling.

Expert bodies, such as the American Cancer Society have considered a family history as an indication for screening or surveillance only for cancers of the breast,

prostate, colorectum, endometrium for reasons of cost-effectiveness and the perceived impact on public health at large [51]. Such guidelines may miss many other familial cancers of moderate risk. Familial cancer risks ascertained by population-based databases suggest that the bulk of cancers may be included in this moderate risk category, which does demand increased clinical surveillance and screening [6]. We have estimated that the American Cancer Society Guidelines on familial colorectal cancer surveillance consider a familial risk of 2.2 an action level and recommend regular clinical screening [6]. Of course, many types of data need to be considered for an action level for familial cancer clusterings, but a familial risk of 2.2 will be exceeded in many cancers, at least in some subgroups.

The level of cancer genetics services varies extensively nationally and internationally. In advanced clinics the counsellor pays attention to the number of affected family members with any kind of cancer as well as the age of onset [52]. The other extreme is that no attention is paid to the family history, although in the past years the level of knowledge on genetic causes for breast, colorectal and some other cancers has reached all modern cancer clinics. However, even if cancer genetics services are in place, there are two other important factors that must be considered [44]. First, the primary health service at the front end of the referral chain, needs to take a family history, to recognise familial cancer clusterings and to communicate the significance of this to the patient [53]. Second, the patients need to understand the significance of an increased familial cancer risk such that an appointment with a cancer genetics professional is made and kept.

The successes in early diagnosis and management of heritable cancers have shown evidence that inherited cancers are no longer inevitable. Although genetic testing has become an important method for at-risk families of certain cancers, it covers a small proportion of all familial cancers. Empirical risk estimates from epidemiological studies have proven to be useful for familial breast and prostate cancers, and it is recommended that reliable risk estimates for other cancers also be seriously considered for routine clinical recommendations. A familial low-penetrant risk in a common cancer may result in more cases in a given population than a high risk of a rare cancer. Furthermore, familial risks are higher for some relatively rare neoplasms, causing hardship in the affected families. Implementation of a unified management plan for familial cancers at large will be a challenge to the involved professionals, the patients and their families and the societies.

## Conclusions

Clinical presentation of a cancer, combined with a family history, may be pathognomonic, diagnostic or suggestive of a syndrome and a cancer genetics professional may readily be able to recommend a gene test in the setting of genetic counselling. However, when syndrome-defining features are lacking, as in the case of mutation-negative familial breast and colorectal cancer clusters, or in low-penetrant cancers, familial aggregation is present, but Mendelian patterns cannot be ascertained, genetic mechanisms have not been worked out and the only types of available risk estimates are derived from epidemiological studies. For some cancers, screening tests are available and they may be recommendable in familial cases, irrespective of whether the genetic background of the disease is known. Familial cancer clustering, without obvious heritability, poses a major challenge to current cancer risk assessment and management. It will be an important task to translate the population-based data on familial risks to the individual-based cancer genetics counselling.

Sometimes familial cancers are dismissed because they are relatively rare among cancers at any sites. The rareness is true, but, in addition to age, family history is probably the only known risk affecting all types of cancer. For some cancers, such as prostate cancer, it is the most prevalent risk factor, after age. There is a further argument for the study of familial cancers. Many of the genes that underlie heritable cancer syndromes are also important in sporadic cancers as somatic events. The gene for Li-Fraumeni syndrome, *p53*, is the most common somatic genetic alteration in cancer, found in half of all cancers. Other examples are the *APC* and *VHL* genes in colorectal and renal cancers, respectively. However, even if this paradigm has not been true for *BRCA1/2*, one of the major motivations for studying familial/heritable cancers is the belief that they will contribute to the understanding of cancers at large. Through uncommon cancers we probably learn more about common cancers.

Finally, epidemiological results on familial risk, familial SIRs, are sometimes dismissed because "they cannot be translated into individual risks". In this article we show several examples on every-day use of epidemiological data in clinical risk estimation. In fact, in the absence of known genes, reliable individual risk estimates need to be generated among populations.

## Acknowledgements

Supported by Deutsche Krebshilfe, the Swedish Cancer Society and the EU, LSHC-CT-2004-503465.

The Family-Cancer Database was created by linking registers maintained at Statistics Sweden and the Swedish Cancer Registry.

## References

- Lynch H, Fusaro R and Lynch J. Hereditary cancer in adults. *Cancer Detect Prev* 1995; 19: 219-233.
- Lichtenstein P, Holm N, Verkasalo P, Iliado A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A and Hemminki K. Environmental and heritable factors in the causation of cancer. *N Engl J Med* 2000; 343: 78-85.
- Hemminki K and Li X. Familial risks of cancer as a guide to gene identification and mode of inheritance. *Int J Cancer* 2004; 110: 291-294.
- Easton D. The inherited component of cancer. *Br Med Bull* 1994; 50: 527-535.
- Vogel F and Motulsky A. *Human Genetics: Problems and Approaches*. Heidelberg: Springer 1996.
- Hemminki K, Li X and Czene K. Familial risk of cancer: data for clinical counseling and cancer genetics. *Int J Cancer* 2004; 108: 109-114.
- Goldgar DE, Easton DF, Cannon-Albright LA and Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994; 86: 1600-1607.
- Hemminki K and Vaittinen P. Familial cancer in Sweden: population-based study. *Int J Oncol* 1997; 11: 273-280.
- Hemminki K and Li X. Familial risk of cancer by site and histopathology. *Int J Cancer* 2003; 103: 105-109.
- Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001; 358: 1389-1399.
- Johns LE and Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol* 2001; 96: 2992-3003.
- Hemminki K, Li X and Czene K. Cancer risks in first-generation immigrants to Sweden. *Int J Cancer* 2002; 99: 218-228.
- Hemminki K and Li X. Cancer risks in second-generation immigrants to Sweden. *Int J Cancer* 2002; 99: 229-237.
- Hemminki K and Li X. Cancer risk in Nordic immigrants and their offspring in Sweden. *Eur J Cancer* 2002; 38: 2428-2434.
- Ioannidis JP, Ntzani EE and Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004; 36: 1312-1318.
- Hemminki K and Czene K. Attributable risks of familial cancer from the Family-Cancer Database. *Cancer Epidemiol Biomarkers Prev* 2002; 11: 1638-1644.
- Lynch HT and de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003; 348: 919-932.
- Aaltonen L, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomäki P, Chadwick R, Kääriäinen H, Eskelinen M, Järvinen H, Mecklin JP and de la Chapelle A. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med* 1998; 338: 1481-1487.
- Salovaara R, Loukola A, Kristo P, Kaariainen H, Ahtola H, Eskelinen M, Harkonen N, Julkunen R, Kangas E, Ojala S, Tulikoura J, Valkamo E, Järvinen H, Mecklin J, Aaltonen L and de la Chapelle A. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 2000; 18: 2193-2200.
- Samowitz WS, Curtin K, Lin HH, Robertson MA, Schaffer D, Nichols M, Grunthal K, Leppert MF and Slattery ML. The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology* 2001; 121: 830-838.
- Hemminki K and Li X. Familial colorectal adenocarcinoma and hereditary nonpolyposis colorectal cancer: a nation-wide epidemiological study from Sweden. *Br J Cancer* 2001; 84: 969-974.
- Olsson L and Lindblom A. Family history of colorectal cancer in a Sweden county. *Fam Cancer* 2003; 2: 87-93.
- Eng C. Genetics of Cowden syndrome: through the looking glass of oncology. *Int J Oncol* 1998; 12: 701-710.
- Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, Loukola A, Bignell G, Warren W, Aminoff M, Hoglund P, Järvinen H, Kristo P, Pelin K, Ridanpää M, Salovaara R, Toro T, Bodmer W, Olschwang S, Olsen AS, Stratton MR, de la Chapelle A and Aaltonen LA. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 1998; 391: 184-187.
- Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR and Cheadle JP. Inherited variants of MYH associated with somatic G:C->T:A mutations in colorectal tumors. *Nat Genet* 2002; 30: 227-232.
- Enholm S, Hienonen T, Suomalainen A, Lipton L, Tomlinson I, Karja V, Eskelinen M, Mecklin JP, Karhu A, Järvinen HJ and Aaltonen LA. Proportion and phenotype of MYH-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. *Am J Pathol* 2003; 163: 827-832.
- Sampson JR, Dolwani S, Jones S, Eccles D, Ellis A, Evans DG, Frayling I, Jordan S, Maher ER, Mak T, Maynard J, Pigatto F, Shaw J and Cheadle JP. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet* 2003; 362: 39-41.
- Fleischmann C, Peto J, Cheadle J, Shah B, Sampson J and Houlston RS. Comprehensive analysis of the contribution of germline MYH variation to early-onset colorectal cancer. *Int J Cancer* 2004; 109: 554-558.
- Gruber SB, Ellis NA, Scott KK, Almog R, Kolachana P, Bonner JD, Kirchhoff T, Tomsho LP, Nafa K, Pierce H, Low M, Satagopan J, Rennert H, Huang H, Greenson JK, Groden J, Rapaport B, Shia J, Johnson S, Gregersen PK, Harris CC, Boyd J and Offit K. BLM heterozygosity and the risk of colorectal cancer. *Science* 2002; 297: 2013.
- Cleary SP, Zhang W, Di Nicola N, Aronson M, Aube J, Steinman A, Haddad R, Redston M, Gallinger S, Narod SA and Gryfe R. Heterozygosity for the BLM (Ash) mutation and cancer risk. *Cancer Res* 2003; 63: 1769-1771.
- Potter J. Colorectal cancer: molecules and populations. *J Natl Cancer Inst* 1999; 91: 916-932.
- Potter J. At the interfaces of epidemiology, genetics and genomics. *Nature Reviews Genet* 2001; 2: 142-147.
- Houlston RS and Tomlinson IP. Polymorphisms and colorectal tumor risk. *Gastroenterology* 2001; 121: 282-301.
- De Jong MM, Nolte IM, te Meerman GJ, van der Graaf WT, de Vries EG, Sijmons RH, Hofstra RM and Kleibeuker JH. Low-penetrance genes and their involvement in colorectal cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 2002; 11: 1332-1352.
- Hemminki K and Chen B. Familial risk for colorectal cancers are mainly due to heritable causes. *Cancer Epidemiol Biomarkers Prev* 2004; 13: 1253-1256.
- Esteve J, Benhamou E and Raymond L. *Statistical Methods in Cancer Research*. IARC Scientific Publication. Lyon: IARC 1994.

37. Hemminki K, Vaittinen P, Dong C and Easton D. Sibling risks in cancer: clues to recessive or X-linked genes? *Br J Cancer* 2001; 84: 388-391.
38. Hemminki K and Li X. Familial colorectal adenocarcinoma from the Swedish Family-Cancer Database. *Int J Cancer* 2001; 94: 743-748.
39. Dong C and Hemminki K. Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. *Int J Cancer* 2001; 91: 144-150.
40. Carstensen B, Soll-Johanning H, Villadsen E, Söndergaard J and Lynge E. Familial aggregation of colorectal cancer in the general population. *Int J Cancer* 1996; 68: 428-435.
41. Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC and Colditz GA. Comparison of risk factors for colon and rectal cancer. *Int J Cancer* 2004; 108: 433-442.
42. Hemminki K and Chen B. Familial risk for colon and rectal cancers. *Int J Cancer* 2004; 111: 809-810.
43. IARC. Weight Control and Physical Activity. IARC Handbooks of Cancer Prevention. Lyon: IARC 2002.
44. Hemminki K and Eng C. Clinical genetic counselling for familial cancers requires reliable data on familial cancer risks and general action plans. *J Med Genet* 2004; 41: 801-807.
45. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C and Mulvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989; 81: 1879-1886.
46. Bondy ML and Newman LA. Breast cancer risk assessment models: applicability to African-American women. *Cancer* 2003; 97: 230-235.
47. Claus EB, Risch NJ and Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* 1990; 131: 961-972.
48. Amir E, Evans DG, Shenton A, Laloo F, Moran A, Boggis C, Wilson M and Howell A. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J Med Genet* 2003; 40: 807-814.
49. Tyrer J, Duffy S and Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004; 23: 1111-1130.
50. Sijmons RH and Burger GT. The use of a diagnostic database in clinical oncogenetics. *Heredit Cancer in Clin Pract* 2003; 1: 31-33.
51. Smith RA, Cokkinides V, von Eschenbach AC, Levin B, Cohen C, Runowicz CD, Sener S, Saslow D and Eyre HJ. American Cancer Society guidelines for the early detection of cancer. *CA Cancer J Clin* 2002; 52: 8-22.
52. Hampel H, Sweet K, Westman J, Offit K and Eng C. Referral for cancer genetics consultation: a review and compilation of risk assessment criteria. *J Med Genet* 2004; 41: 81-91.
53. Sweet KM, Willis SK, Ashida S and Westman JA. Use of fear-appeal techniques in the design of tailored cancer risk communication messages: implications for healthcare providers. *J Clin Oncol* 2003; 21: 3375-3376.